MHP
A PORSCHE COMPANY

A.I.dea Book

Use Cases & Trends in Artificial Intelligence

EDITORIAL

Artificial Intelligence (AI) at MHP is a close-knit, cross-functional global community where consultants from almost all organizational units' join forces to harness synergies between Machine Learning, Data Science, Software Engineering and Business Domain Knowledge. This internal setup pays tribute to the cross-functional and interdisciplinary nature of AI, which on the customer-side means optimally staffed projects. Members of our AI Community actively participate not only in the delivery of projects or the implementation of products but also to exchange knowledge, experience, and practical insights in the AI domains Machine Learning, Computer Vision and Natural Language Processing.

Early this year in one of our vibrant AI Community Meetings, we compared notes on issues customers are currently facing concerning the utilization of data-driven technologies. Interestingly, we soon found that our consultants corroborated findings reported in recent studies, stating that companies have recognized AI's potential. By the same token the expectation that by 2030 some 70% of companies will have adopted at least one type of AI technology does not seem far-fetched. However, despite these clear intentions, often driven by the fear of missing the "technological cue," we agreed that many of us detected signs of what can only be described as being haphazard AI-campaigns often lacking even a clearly identifiable use case.

We asked ourselves why AI is evidently having traction problems in delivering on its promise, and found that companies are having difficulties transferring technical principals to viable use cases. Furthermore, companies that were able to overcome initial obstacles and identify suitable use cases had often gone a step further and successfully implemented proofs-of-concept, which in approximately 80% of all cases failed in the critical next step: scaling to the production stage and integrating AI components into existing system landscapes.

This A.I.deaBook – essentially a documentation of our AI-related Ideas, is the brainchild of this AI-Community Meeting of early 2019. We decided that it was time to put our thoughts on paper, to write down our observations from the frontier and provide poignant practical experiences to which our customers can relate. We think that this series of articles will prove to be beneficial for all those interested in making AI work to create business value by providing insights into use cases for which we have implemented solutions and by sharing our experiences in scaling these to production.

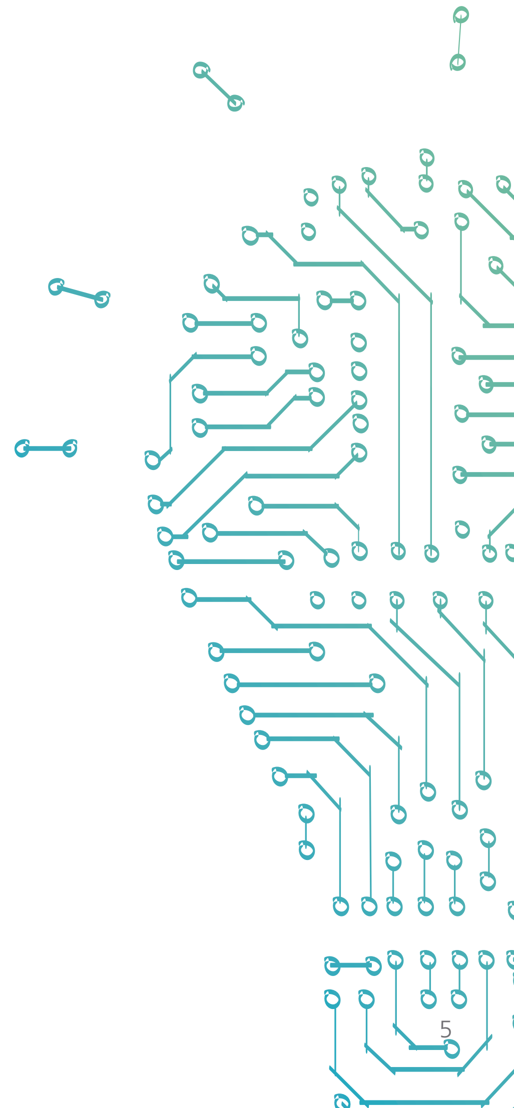
We appreciate that successful full-scale AI solutions produce measurable improvements in efficiency and/or quality in their specific areas of application. On the technological side, this is best achieved by applying technology, which enables rapidly scalable AI solutions based on a containerized approach – We call this **AI as a service**. It is our firm belief that the adoption of this approach will be decisive over the success or failure of full-scale data-driven business applications in the foreseeable future.
On the people-side, solid technological foundations must be backed by the right methodology while utilizing the gained business advantages responsibly.

To highlight our thoughts in the areas of efficiency, quality, technology, methodology, and responsibility, MHPs A.I.deaBook presents 17 articles in these thought-dimensions written by more than 25 MHP consultants, demonstrating how we help our customers to shape a better tomorrow.

Johannes Keller &

William Cobbah

TABLE OF CONTENTS



QUALITY

01

Automated Visual Inspection

By Fanli Lin and Holger Muth-Hellebrandt

A Small Step for AI, but a Huge Leap for Manufacturing

A rising tide of automation and intelligence is sweeping across the globe. To improve production efficiency and increase overall profits, many industrial companies have adopted various technologies and solutions in their manufacturing processes. The buzzword here is "Smart Manufacturing" or "Industry 4.0", which refers to the improvements in manufacturing through the implementation of different technological fields like "Artificial Intelligence", "Robotics", "Industrial Internet of Things" and "Cyber Security". The Integrated Factory System (IFS) launched by MHP simulates the smart factory of tomorrow and serves as a live example for visitors to experience and "physically touch" Industry 4.0. It contains various so-called "Industry 4.0 technologies", such as an Automatic Guided Vehicle System, Voice Control, SAP Extended Warehouse Management, or Virtual Reality. Most recently, it is extended by an AI-based visual inspection station, a detailed explanation of which will be the focus of this article. Figure 1 shows the overall setup of the integrated factory system and where our visual inspection (VI) station is located.

Automatic Visual Inspection in Industry

In a highly competitive market, companies need quality and productivity. Especially in manufacturing, many repetitive quality control tasks are performed as manual labor. However, the visual inspection performed by humans is often time-consuming and inaccurate regarding the inspection quality. Human errors caused by environmental and personal factors ultimately result in higher costs. Automatic visual inspection solutions enable companies to increase the throughput of their assembly line while reducing operating costs and improving product quality.

Our Three-Step Approach

To simulate automatic visual inspection in a real manufacturing process, we split the tasks into three steps:

1. Object Detection: Identify different loads/objects on a conveyor belt in real-time using image processing techniques with OpenCV. Here advanced methods like deep learning are unnecessary because the traditional computer vision techniques deliver very good results with little resources.

2. Anomaly Detection: Once the target object has been detected, the software must detect whether there is an anomaly or defect on the surface of the object. This step is usually where the actual quality control takes place. For the purpose of our showcase, we defined our anomaly as any mathematical expression written by humans.

3. Anomaly Interpretation: On our clients' assembly lines, there is often an extra process after anomalies have been detected. In these processes, anomaly types are get-

Figure 1 shows the overall setup of the integrated factory system and where our visual inspection (VI) station is located ting categorized or even interpreted for subsequent operations. We simulated this real-world process by interpreting the mathematical expression on the surface, our anomaly, and calculating the result.

Object Detection

As opposed to a data-driven approach in deep learning where the characteristics of an object are learned by training a model, our approach is "algorithm-driven", which means that we manually design an algorithm that can detect the typical characteristics of our target object using image processing techniques. For our object, the major characteristics are the white, gray, blue colors and the two squares in different sizes (see Figure 2). If the color and the shape of an object meet the requirements, our implementation declares an object as "detected".

For image processing, we used OpenCV[1], which is a very powerful open-source computer vision library, and contains more than 2500 optimized algorithms. The image below illustrates the major steps in our algorithm:

1. Preprocessing (Resize, Color Space Conversion, Image Blurring, and Channel Splitting)
2. Thresholding to get the binarized image
3. Find the inner square
4. Find the outer square
5. Target is detected if both squares are found

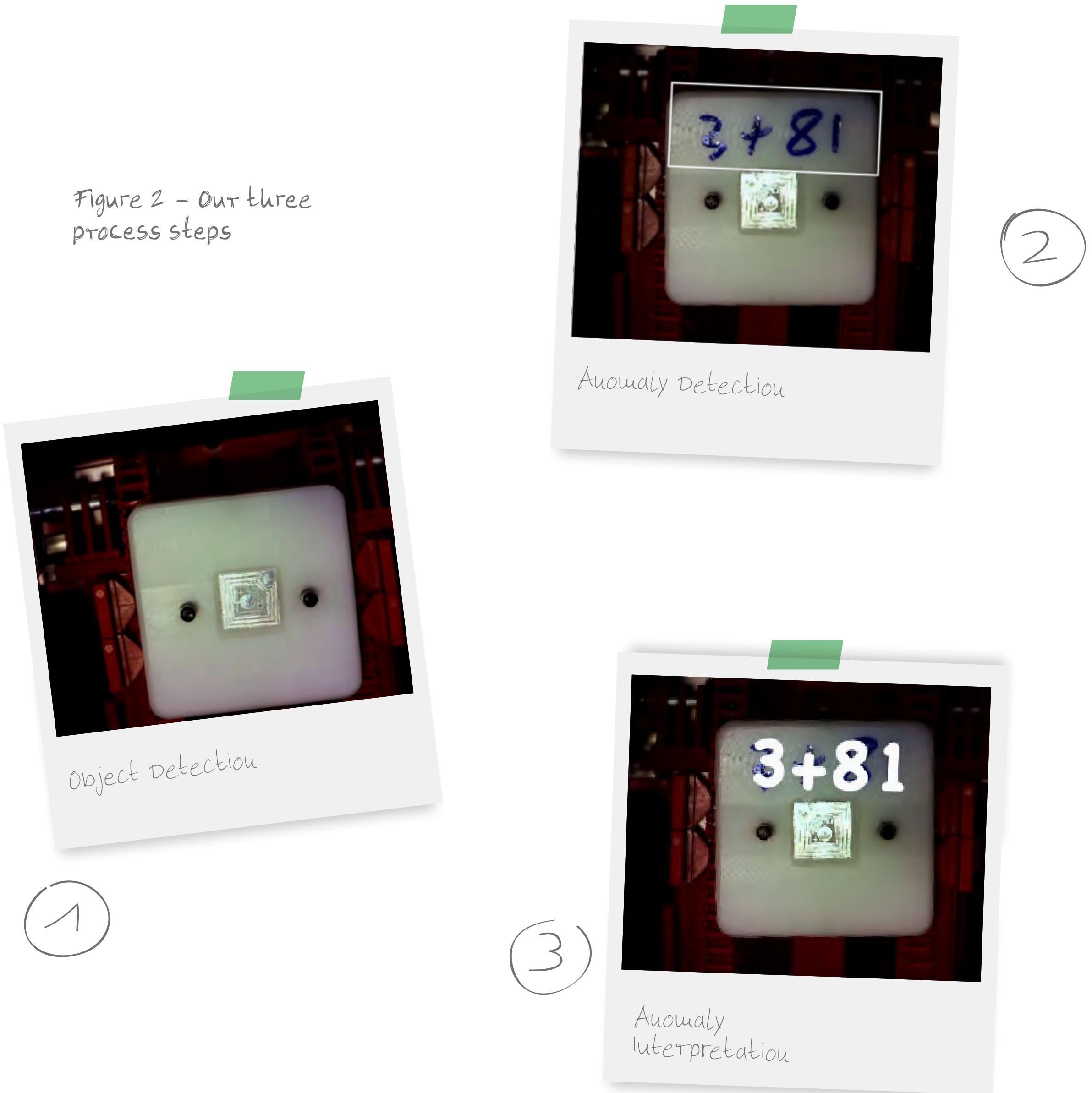
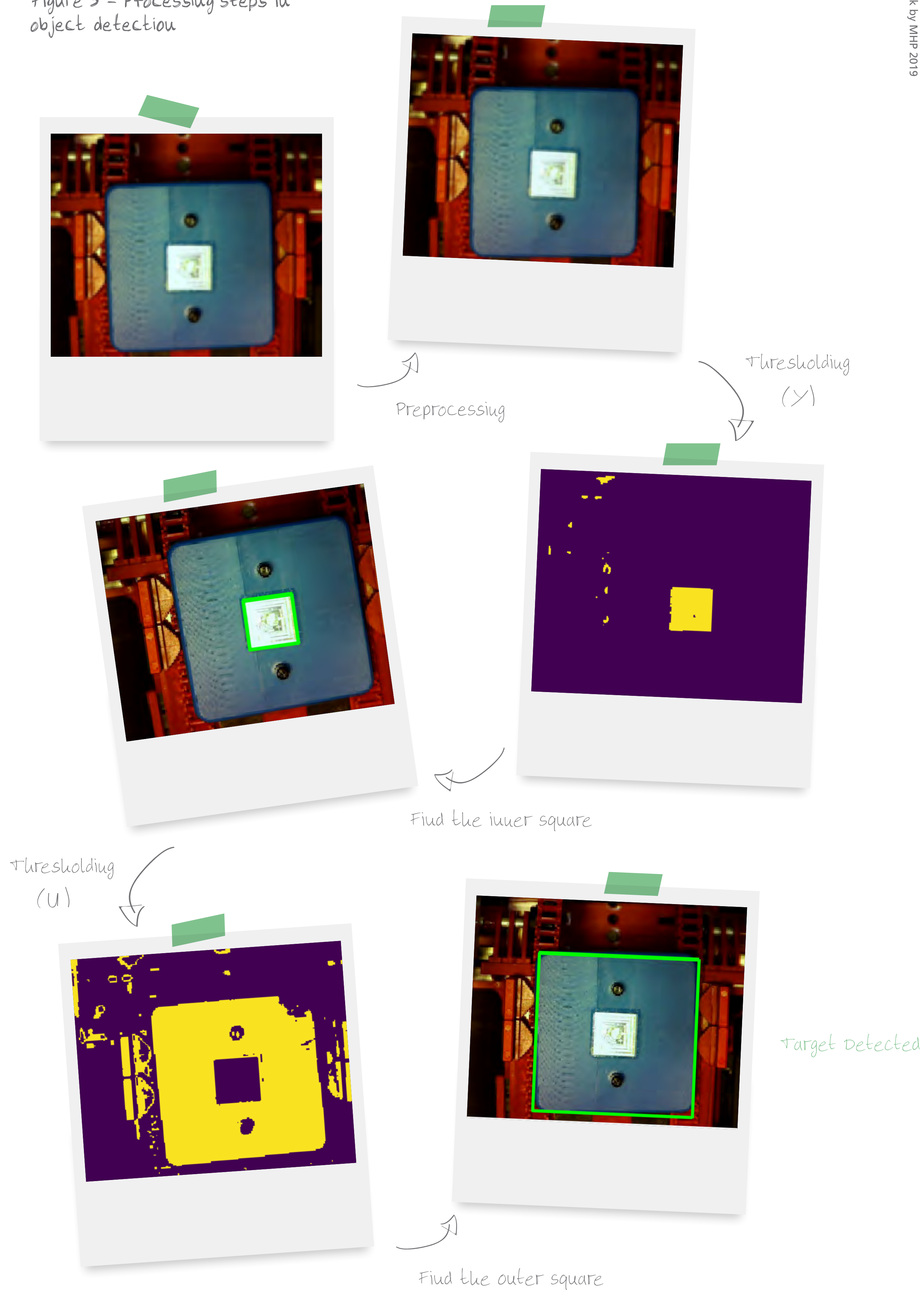


Figure 2 – Our three process steps

1) For more information on OpenCV visit https://opencv.org/about/

Figure 3 – Processing steps in object detection



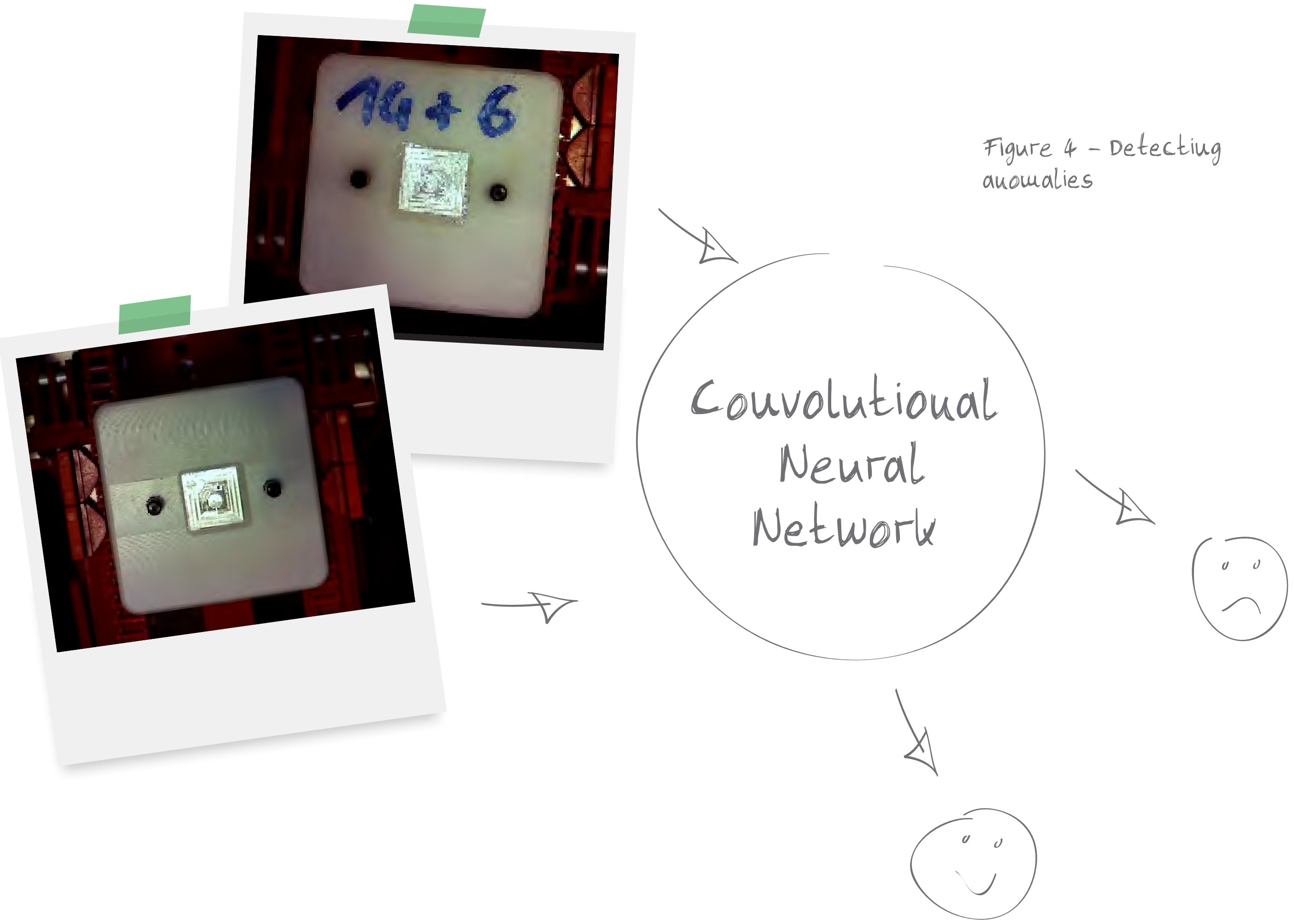


Figure 4 – Detecting anomalies

Anomaly Detection

According to Chandola [1], Anomaly Detection is defined as the problem of finding patterns in data that do not conform to expected behavior. Anomaly detection has many applications in various domains, e.g., medical diagnosis, industrial quality control, and fraud detection. In Computer Vision, one needs to differentiate two scenarios in anomaly detection [2]. The first one is the classification scenario, where anomalies appear as entire images, which should be labeled as outliers. The second scenario is the segmentation scenario, where anomalies manifest themselves in normal patterns, and localizing them in images is required. In our case, we focus on the first scenario where we classify an image as "normal" or "abnormal". To be more specific, we define our anomaly detection as a binary classification problem.

The technique we used here is known as "transfer learning", where we took a pre-trained neural network and trained an extra layer on top of the network to differentiate the normal image from the images with anomalies as shown in Figure 4. Our testing result shows that our model has an accuracy over 99% and predicts the result in less than 0.001 seconds.

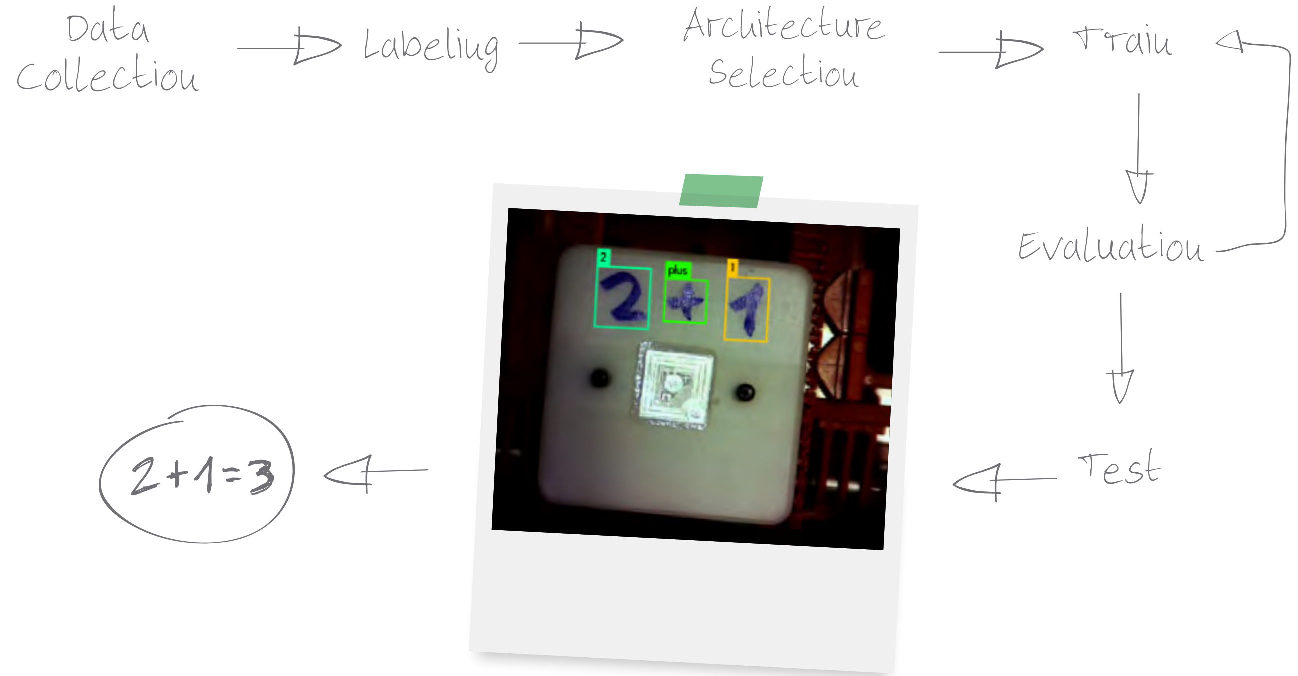
Anomaly Interpretation

The aim of this step in our example is to recognize a handwritten mathematical expression and calculate the result, which is slightly different from the real-life scenario where different types of anomalies are classified based on whole images using multi-class classification techniques. Our task here is more like Handwritten Digits Classification in the field of Optical Character Recognition (OCR) but with additional classes for the mathematical standard operations addition, subtraction, multiplication, and division. One approach in the pattern recognition community is to first apply the Histograms of Oriented Gradients (HOG) image descriptor to extract features from the digital images and then use a multi-linear Support Vector Machine (SVM) for the classification [3]. The drawback compared to a deep learning approach is that it is not an end-to-end system that recognizes characters directly from raw pixel data, which means that it needs to find the digits first using image processing techniques and then apply a classifier.

Therefore, we applied the popular object detection framework YOLO[2] ("You only look once"). YOLO lets us train a deep neural network that can detect digits and operators directly from the raw image. The whole pipeline contains data collection, labeling, network architecture selection, training, evaluation and testing as illustrated in Figure 5.

2) For more information on YOLO models visit https://arxiv.org/abs/1506.02640

Figure 5 – Sequence of the YOLO Pipeline



Architecture

The idea of the underlying backend architecture is to set scalability of our VI as a service solution as our number one priority to meet real-world industrial requirements. The VI service is, therefore, hosted on public cloud – as a Docker container orchestrated as part of a Kubernetes cluster. The benefits of using such a cluster architecture are as follows:

- **Vertical & Horizontal Auto-Scaling**
Kubernetes offers the ability to auto-scale a service vertically (scaling the number of running service instances) as well as horizontally (scaling the resource requirements of running instances) by defining groups of containers with shared storage/network known as "pods" as the smallest deployable units [4] running on nodes inside a cluster. That way, the architecture provides the user with full resource control but also with the ability to scale the hosted services as required.

- **Loose Coupling**
The deployed solution in the cluster runs with load balancers as a buffer between services and instances. Running such a setup makes it easier to scale services independently. Increasing load on one service can be handled without adjusting or reconfiguring any of the other components and layers of the cluster, which may be hosting other components of the software.

- **High Availability**
A failing service remains available as Kubernetes uses liveliness and readiness parameters permanently on all instances of a service. With this permanent availability control, failing service instances are restarted and redeployed immediately. Until they are ready again, they will not become targets of the load-balancer's request dispatching.

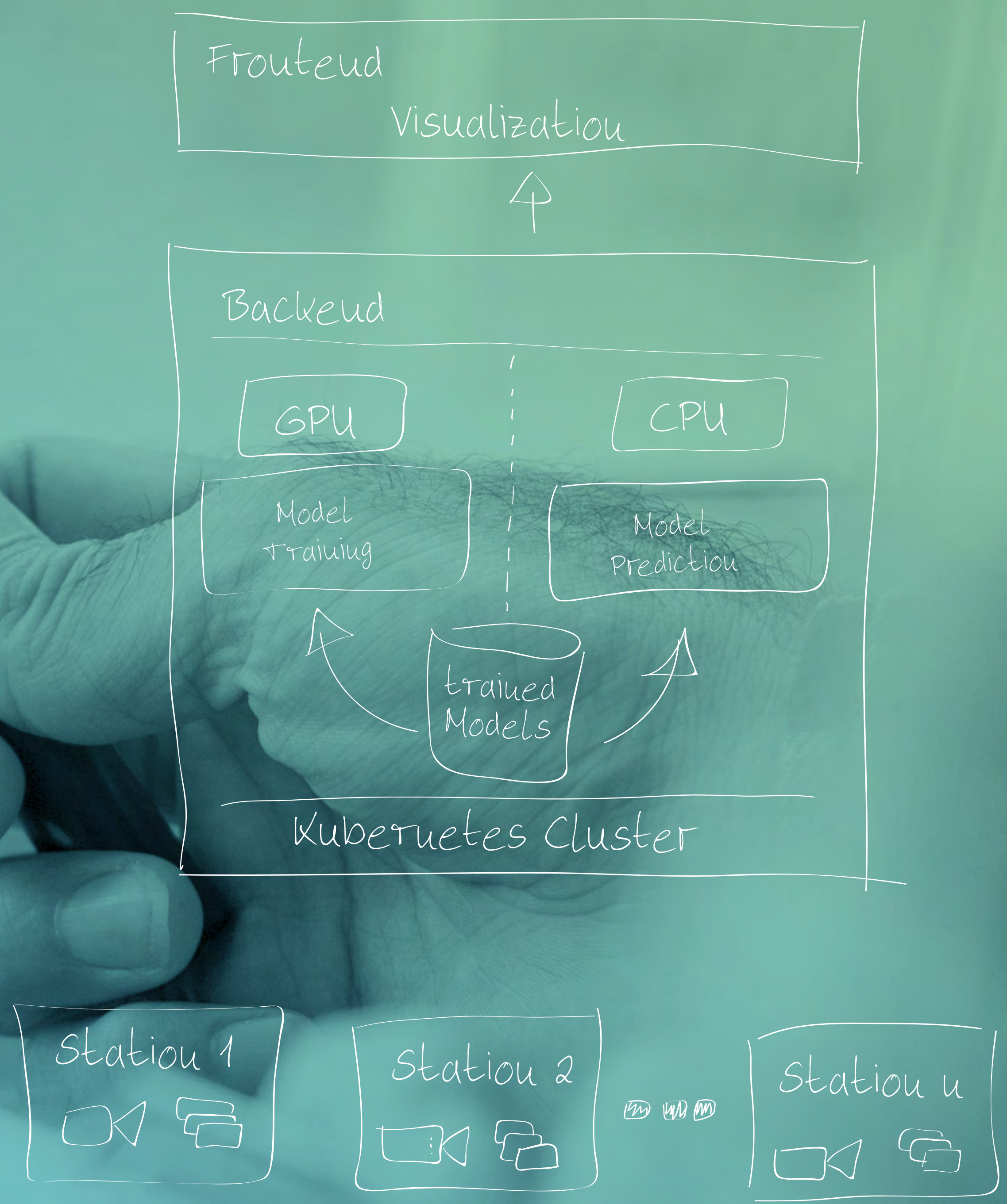
So, resources in the backend can be scaled up or down according to the number of requests and traffic. Additionally, the system can also be easily scaled up with regards to the number of edge-devices as this is one centrally run cloud-service endpoint (see Figure 5).

Also, the solution is ready to be extended by enterprise features, e.g. a notification service, logging service, etc. To summarize; the architecture concept, using a Kubernetes cluster as a service platform leads to more reusable modularity of the deployed services as well as cost reduction. Regarding the VI showcase and its three major steps (Object Detection, Anomaly Detection, and Anomaly Interpretation), two services are hosted on CPU-based cloud instances: one service hosting the Tensorflow model to predict handwritten anomalies and one for the YOLO model to interpret the detected anomalies. In addition, the nodes in the cluster can be scaled up for the training of the Deep Learning Model as this requires GPU-based instances.

In a nutshell, we apply image processing techniques to recognize different loads/objects on the conveyor belt, then used deep learning models to find and interpret the anomalies on the surface. The whole system is deployed as multiple microservices in containers on a Kubernetes cluster to enable scalability and reusability.

We have successfully implemented this solution for multiple clients. The great benefit of our approach is that it easily integrates with or extends existing assembly line procedures.

Figure 6 – Visual Inspection architecture



Funneling Fakes

By Beatriche Gencheva, Arno Fuhrmann

A First Line of Defense in Online Fraud

Business related fraud has reached an alarming scale and presents a significant problem for a large number of businesses worldwide. Although many resources are being invested to recognize and prevent it, fraud detection is not a simple task and requires complex and time-consuming investigations that deal with different knowledge domains. The fact that Facebook detected and removed 1.7 billion fake accounts from its platform between July and September 2019 shows the dimension of the problem. Currently, the number of Facebook accounts with suspected fraudulent activities is approaching 5.4 billion, with 3.3 billion deleted fake accounts in 2018. [1]

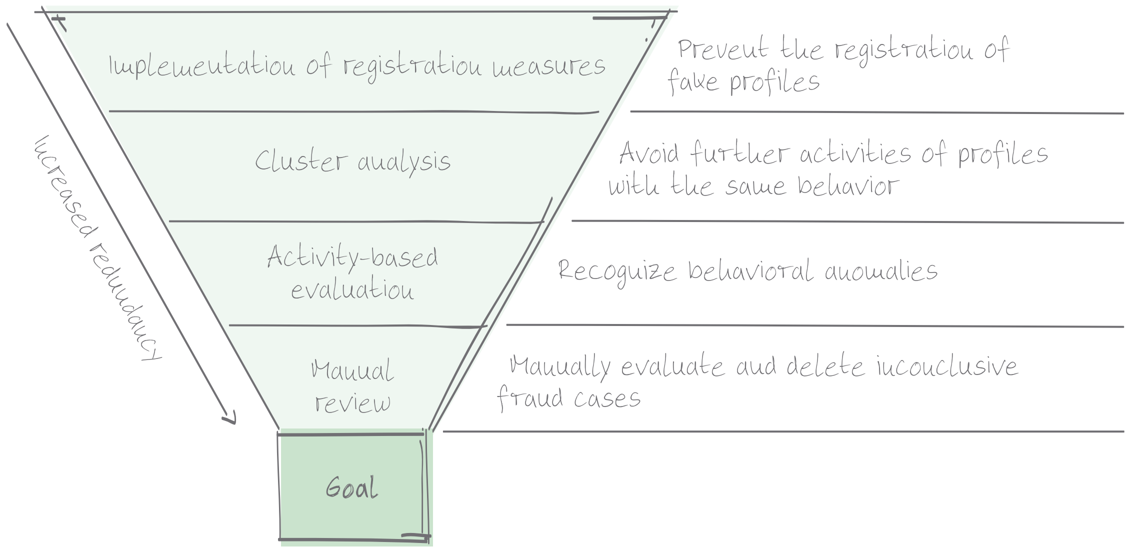
In a data-driven world of ever-increasing complexity, the possibilities to commit fraud are immense, and the intentions of fraudsters are diverse, limited only by their personal "creativity". Until recently, the buzzword "fake account" was primarily associated with the adverse influence of user behavior in online social networks. However, the malicious, large-scale creation of fake accounts is expanding its focus and is already causing severe financial losses to various companies, especially those whose business models center on monetarizing customer relationships. Detecting these types of attacks is high on the agenda for affected platforms and businesses, but since fraud is an adaptive crime, unique methods of intelligent data analysis are required for detection and prevention to provide a significant advantage for crime-fighters in this ever-ongoing game of cat-and-mouse.

Currently, many organizations are using rule-based approaches to identify fraudulent activities. Although rules can be very accurate in uncovering already known patterns, they face substantial challenges when it comes to the recognition of possible relationships and new fraud patterns. Modern techniques are transforming the vast amounts of collected data in ways that enable the intelligent linkage of data, which in turn, allows the inference of conclusions, ultimately leading to the ability to predict or detect fraud.

From MHP's perspective, the detection of fake online accounts is yet another AI journey, which we choose to embark upon with our customers, requiring a systematic approach to break down complexity. For these purposes, we suggest a "funneled" approach consisting essentially of multiple lines of defense, each aiming to reduce the data volume by eliminating accounts/data from further processing, which by virtue of being real accounts, withstand the scrutiny of preceding activity layers of the funnel (see Figure 1). Since machine learning provides algorithms to identify patterns in new data, it is ideally suited for detecting and predicting ever more adaptive methods of online fraud and is therefore used in different variations in several stages of the funnel.

A similar approach was introduced by LinkedIn [2] to manage the surge in fake accounts in online social networks, which aim to deceive or disturb real users. As opposed to this approach, which incorporates reports of suspicious activity, our funnel defense solution includes additional methods for activity detection, which consider the initial actions of potential fraudsters making it more generically applicable outside the realms of social networking.

Figure 1 – Funnel Defense Approach for Detection of Fake Profiles



The first stage aims to prevent the creation of fake accounts at the very point of registration. More-or-less simple data validation routines (plausibility checks) are used to verify data at the point of data entry. Additionally, several different tests may be employed in order to differentiate between human and computer interaction. These tests are not limited simply to data entry but include behavioral biometrics, which utilize machine learning algorithms to identify abnormalities in user behavior during registration.

Although quite trivial in terms of test complexity, this first stage in our funnel serves to prevent the creation of a reasonably constant "mass" of fake accounts, together with a more variable volume of accounts depending on the complexity of the implemented behavioral biometrics. Despite this, a remaining number of accounts being created will surpass this first layer due to the algorithms having insufficient features to be able to distinguish between malicious and real profiles.

The second stage of the funnel clusters created accounts in groups in order to find abnormalities from the received data from the first stage. By doing so, it is possible to obtain an indication as to whether profiles were created from the same person or device. The machine learning techniques used are those of supervised learning based on shared attributes in multiple dimensions to infer a common source or origin.

The third stage of our defense-funnel serves the purpose of identifying fake accounts, which adhered to the input rules of stage one and could not be attributed to a common point of origin. Activity-based evaluation is used to observe behavior within the accounts/profiles based on events occurring after new profiles have been generated. Once again, we employ powerful clustering models in search for new patterns of anomalous behavior. In addition to this, the third layer includes other supervised methods seeking for already known specific types of fraudulent behavior previously identified for similar accounts. The main advantage of the AI@MHP funnel approach is that it breaks down the complexity of the complete analysis into several stages, reducing the volume of data from stage to stage such that residual unclear cases of possible fraud can be identified in the final manual evaluation stage.

There are numerous reasons for machine learning being an ideal extension of existing rule-based mechanisms for the detection of fake profiles. Machine learning methods can deliver performance in terms of speed required for the detection of fraud of this type in real-time and with the increasing amount of data, machine learning systems consistently improve their performance. While being highly cost-effective, the new AI-based solution is more efficient and accurate when it comes to uncovering non-intuitive patterns or new trends, which might not be evident to a human fraud analyst.

Quality Management of Complex Systems

By Jan Janetzko

How AI closes the Loop in Complex Systems' Quality Management

The quality of products we customers buy defines how satisfied we are for the most parts. But quality is not only essential from a customer viewpoint. "Do it right the first time", which was claimed by Edward Deming, who is well known for the PDCA cycle of continuous improvement, was a milestone in quality management. Low scrap rates and high first-pass yield drive productivity and, thereby, a company's profitability.

But the history of quality management dates back even further and has come a long way since. In the early 20th century, when Henry Ford introduced a quality inspector for his famous Model T assembly line. It got big in Japanese companies where resources were scarce after WWII and scrap could not be afforded. And it was only in the 1990s that quality management got standardized on an international scale with the introduction of ISO's 9000 series.

Currently, the optimization of products that have already been introduced into the market mainly comes from either R&D driven experiments or simulation-based physical models. Expert knowledge or information from the field plays a huge role. Products can also be optimized via continuous improvement processes based on the Japanese philosophy of KAIZEN in lean management, where knowledge from production workers is fed back into the product and process design.

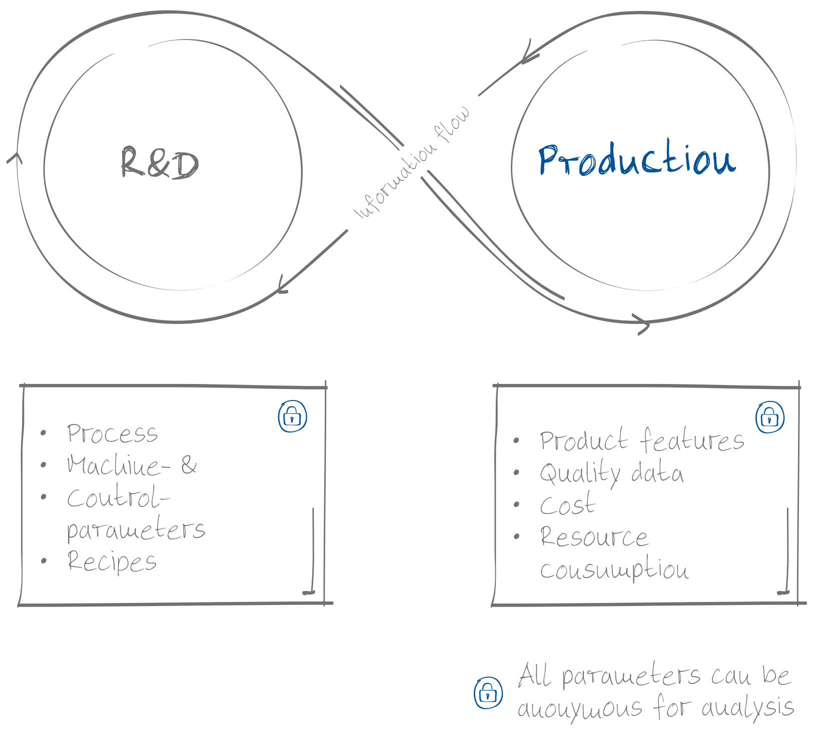
In ever-increasing product complexity and production automatization, these sources of knowledge become more and more difficult to exploit. The challenges in obtaining the necessary information are immense, and there are simply too many parameters and variables to consider. Based on experience, MHP has come to believe that KAIZEN style product optimization in production is becoming increasingly difficult: With an increasing degree of automation, the vast majority of workers no longer fully understand the production process. Experiment- and simulation-based product enhancements are getting increasingly expensive and time-consuming; the more complex a system gets. Engineers are therefore forced to define a level of accuracy, which satisfies the purpose of the simulation. In this case, this step is of great importance because the more precise a simulation gets, the better all parameters and their interactions within the system must be known. In rather simple systems, a high level of accuracy can be obtained with relatively little effort. However, in complex systems, as described in the case studies below, simulations with a high level of accuracy can easily cost millions of euros, and a small mistake in a parameter estimation or their expected interaction will change the simulation result by a magnitude of several orders.

But with the ever-increasing availability of data generated by sensors within connected and fully integrated machines within the production process, completely new possibilities to generate information and knowledge about products from inside production arise. In production processes for automotive components, more than 2500 datasets are needed for process-, machine-, control- and recipe-parameters. The actual values of the parameters with which a product is being produced are monitored, stored, and tracked within the business logic at each production step for traceability reasons (see Figure 10). However, only best-in-class companies go a step further! In most companies, these often locally held databases have a thick layer of digital dust on them. This dust can be blown away and years of historical datasets can be transformed into a mine of pure gold by applying the following idea: Data from along the complete supply chain is stored centrally (which is called a Digital Twin of the product) and then used to feed an Artificial Intelligence (AI) system or application.

The first step is to label the data of the Digital Twin with information from quality management, production control, resource consumption, and data from other product features. Using information on traceability data applied to AI in this manner is termed "labeling." Next, data streams which can contain features from a few hundred to hundreds of thousands of products are fed into a Machine Learning Algorithm, often a Deep Neural Network. The neural networks iteratively weighs correlations between production parameters and the label data from quality management, production control, etc. (see Figure 11). Based on this information, the algorithm evaluates ideal ranges for the production input parameters. Insights gained in this manner are used by domain engineers to reevaluate the specification limits for the process-, machine-, control- and recipe-parameters. After closing the loop by updating the production line accordingly, overall production costs are potentially significantly reduced with a rapid payoff.

This concept arose in recent years within the process industries due to the dramatic advances in AI technology. Here, levels of automation have been very high for at least the last two decades, such that data has been present for a long time. Also, continuous production processes such as in pharmaceutical plants, oil refineries, or steel mills made it easier to maintain all process-related data in closed IT systems that are tailored to the process. With the fourth industrial revolution and the age of the IoT in discrete manufacturing, the areas of application multiply due to data availability and connected IT systems. In the following section three case studies will illustrate how the concept described above can offer substantial value-added in product quality, process quality, and conventional quality control.

Figure 1 – Information flow cycle between R&D and production



Case 1 – Product Quality

AI-based analysis of interactions can enhance product quality within a complex tolerance chain in dynamic environments such as engines, turbines, or transmissions. In terms of quality assurance, the quality of the overall system cannot be equated with the sum of the quality-assured individual parts. Even if all individual parts function perfectly according to their specification, there is no guarantee that the overall system will work faultlessly - this results in a clear benefit by using AI to identify and eliminate sources of error in historical data from production processes and the product quality data labels.

Case 2 – Quality Control

Another use case for AI-driven quality management is to enhance traditional quality management as introduced at Ford at the beginning of the 19th century. A feedback loop is implemented to forecast the probability that a part will become scrap in an early stage of production. Therefore, costs are reduced by not processing the part further or by defining precise rework processes if this probability is above a defined threshold. An example of this is in car body assembly lines, where the quality of welding points is analyzed within the process. The quality of the finished item is then predicted by an algorithm, which was trained with the information from the analyzed data that has been gathered from thousands of previously assembled car bodies.

Case 3 – Process Quality

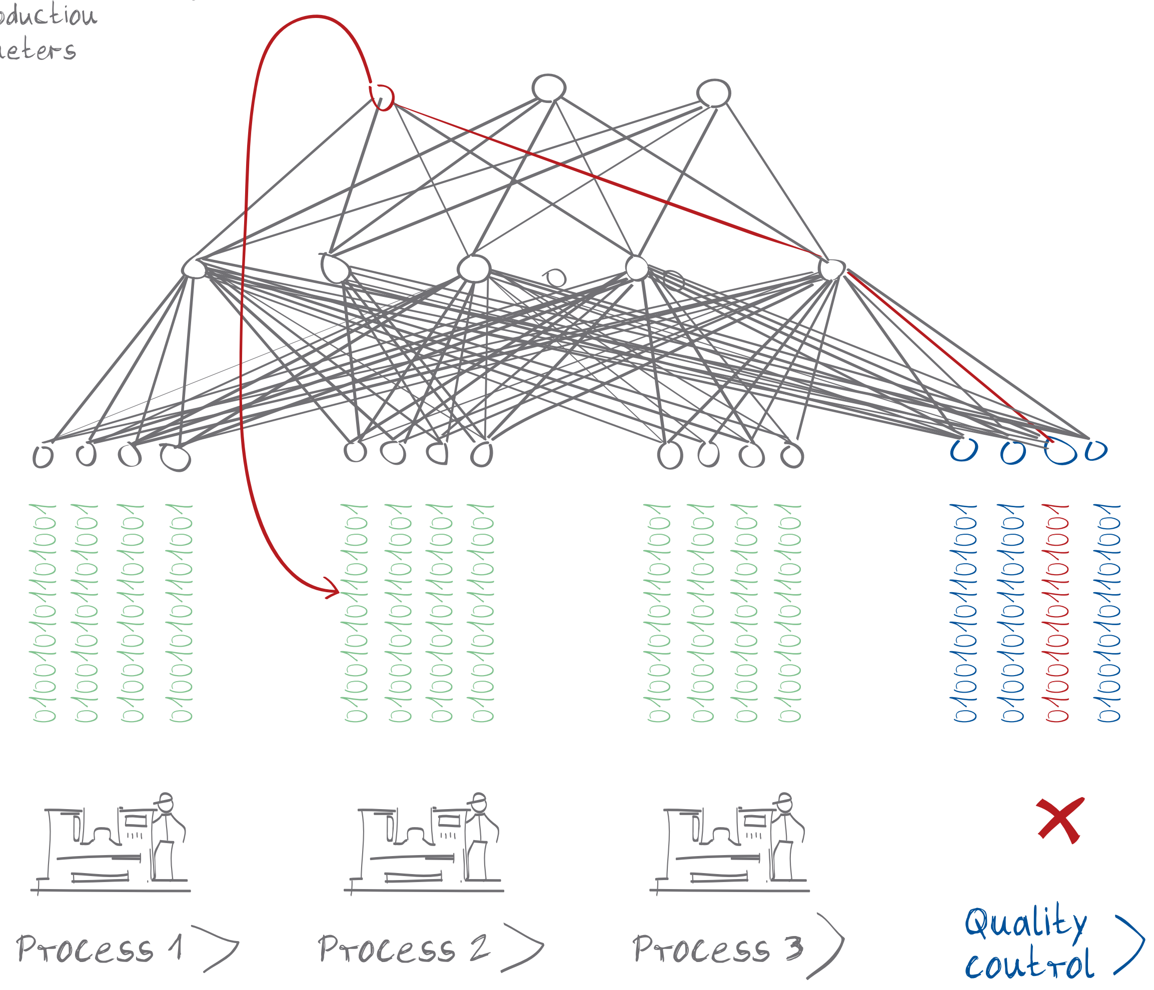
AI can improve not only product quality but also process quality – another pillar of ISO 9000's philosophy. Applied to the optimization of complex, iterative production processes for laser systems, AI supports production engineers in transforming the adjustment process from a highly iterative and heuristic-based process into an algorithmic process. The starting point is an assembled laser, in which all components function perfectly on their own. However, the emergent system properties (laser light, laser power, etc.), only emerge through the interaction of the perfectly aligned subsystems. The more precisely a laser is adjusted, the better the output parameters such as power and beam quality get. In modern lasers, not only two mirrors must be aligned parallel to each other, but a large number of optics must be positioned precisely in the beams' path. As a further challenge, the device does not remain static, as the design suggests. Various thermal and optical effects lead to a time-varying behavior of the system. In an AI-driven alignment process, the mirror positions are labeled with the quality parameters output power and beam quality.

All three case studies show that in quality management of complex products, one plus one rarely equals two. The behavior of a complete system is rarely equal to the behavior of all individual sub-systems or components. For quality management experts, this is a well-known fact. We humans have reached our cognitive limits in fully understanding products and processes and, therefore, our ability to optimize these systems. Precise simulations can be costly and time-consuming. We rely on much simpler heuristics as and input for our optimization efforts and simply assume that the sum of the behavior of subsystems will equal the behavior of the whole system. But new technologies are arising and are enabling industries to overcome these shortcomings as some first movers already demonstrate.

MHP helps its customers in the automotive and manufacturing industries to significantly increase product quality and first-pass yield. As a result, our customers report lower overall production costs by an increase in effective productivity of around 2-3 % (cases 1 and 2) measured in terms of Overall Equipment Effectiveness in already traditionally optimized production and quality management systems. For case 3 this increased productivity can be even a double-digit percentage increase depending on the base line.

Our experts have proven the potential of MHP's closed-loop approach for quality management: It does not only affect the way product and process optimization works but that it also has the potential to revolutionize quality management, by generating deep insights from data that cannot be obtained using the ordinary tools of ISO 9001 style quality management. We are living in the age of omnipresent data – let's use this to our advantage!

Figure 2 – Clustering of production parameters



EFFICIENCY

02

Automation of Production Planning Enhanced by AI

By Till Giese (Porsche Consulting), Steffen Franz, William Cobbah

The Next Level for Pharma Companies

Production Planning as Key Success Factor for an Efficient Production

Lack of quality in planning and scheduling can reduce productivity as measured in Overall Equipment Effectiveness (OEE) by 5%. In order to improve supply reliability and uphold stable production, inventories are often increased [1]. Thus, an excellent planning and scheduling is a significant competitive advantage. In order to achieve an excellent planning and scheduling, pharmaceutical companies have to manage 5 major challenges:

1. Intr-day adjustments in the production sequence
Short term re-sequencing within a frozen zone leads to low adherence to the original production plan. Production is frequently asked to adjust to changes in demand. In order to achieve this flexibility, the inventory is increased causing high levels of throughput times.

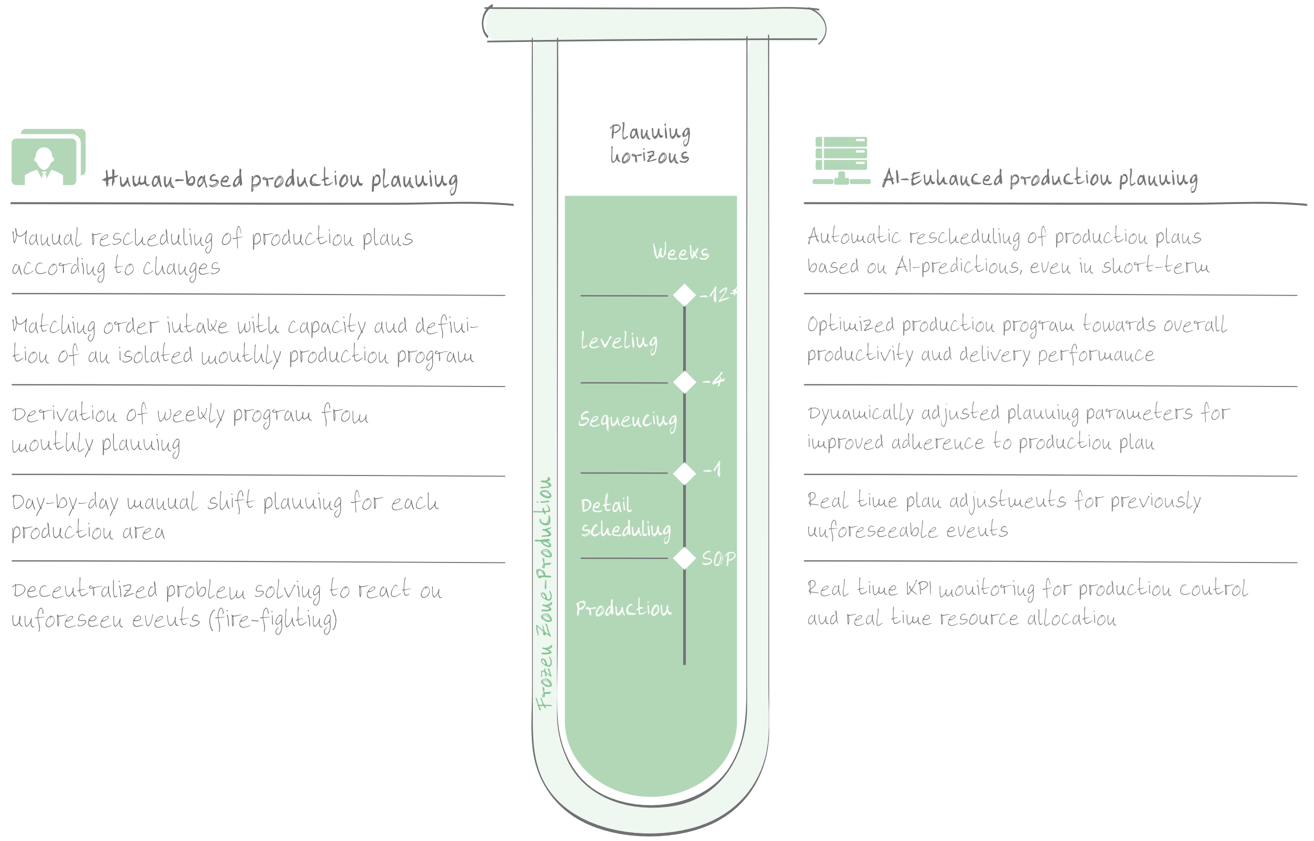
2. Unclear effects of breakdowns on delivery times
Compared to other industries pharmaceutical companies have low OEE, especially in packaging. High product variability and low batch volumes contribute to this. However, equipment breakdowns or reduced efficiency, due to small stoppages, are also contributing factors to a low OEE. As a result of this, lead times become ambiguous, resulting in reduced delivery performance and low customer satisfaction.

3. Reduced OEE due to non-optimal setup sequence
In the pharmaceutical industry, there is a huge variability in changeover times, especially in packaging. If you are switching from one product to another, which is the same drug but just has more tablets per box, just a partial changeover is sufficient. On the other hand under most other circumstances however, a full changeover needs to be done. By following a specific production sequence, the OEE can be improved by 5%. A strong focus on meeting delivery requirements (challenge 1) results in a low focus on the right production sequence.

4. Lack of integration of interface areas
Several interfaces have to be considered in the pharmaceutical value stream, e.g. from bulk to packaging and from packaging to quality control. Planning and scheduling should consider the complete value stream and thereby optimizing the inventory at the interfaces. Especially the interface to quality control is challenging, as capacities in quality control are not reflected by most planning and scheduling tools.

5. Daily fire-fighting activities
Daily troubles challenges, like such as missing employees, equipment breakdowns (challenge 2), and missing material need to be handled. This is quite laborious as this often involves adjusting the production schedule. For example, if employees are missing because due to illness of sickness, the scheduler needs to find the employees with the right set of skills to replace the missing employees. Eventually the scheduling has to be adjusted, because some equipment has

Figure 1 – Comparison of human-based and AI-enhanced production planning



to be at least partially shut down. The challenge is to find the best solution, which has lowest negative effect on OEE, and delivery performance whilst maintaining optimal costs. Since this is very complex, a significant number of working hours is spent on this, causing high labour costs.

Target Picture of an Ideal Production Planning System Enhanced by AI

Challenges as described in section 2 serve as input for the functional specification of an optimal production-planning environment, which may be further broken down by regarding various planning horizons within the production planning process. Figure 1 shows a comparison of production planning approaches, today and tomorrow across short- and long-term planning horizons

Our experience shows that much of today's production planning becomes an increasingly manual task with decreasing planning horizon. This stands to reason since the criticality of changes made to the production schedule in the short-term increases whilst requiring the integration of a much greater number of variables. Traditionally, computers would not have been entrusted with tasks at this level of required "intelligence" and indeed, standard procedural implementations of the logic required to perform alterations to a production schedule within the frozen zone would be neither reliable nor economically viable. However, as is currently observable in many other business scenarios, the introduction of AI in the context of short term production planning becomes a game changer.

Whereas in the vast majority of today's production planning processes, detailed scheduling relies on a manual shift planning for each production area on a day-by-day basis, tomorrow's fully automated scheduling systems will utilize AI to incorporate short term events to generate revised production schedules at equipment and worker level in real time. Looking at the different planning horizons in Figure 1 it becomes clear that the benefits of integrating AI into the planning process increase the close one gets to the start of production. Especially when it comes to short-term adjustments within the frozen zone, due to sudden staffing shortages or reduced machine capacities, an appropriately trained AI-based system can suggest staffing and parameters for the on-board optimizer, regarded as being an optimal solution to the problem since similar patterns are known from previous situations.

One of the major challenges in the operation of mission critical AI-applications is that of transparency when it comes to the inner workings of the AI-model. Despite the fact that methods to understand "black-box predictions" exist [3], these are complex such that a more practical approach is required. Modern Planning Schedulers allow interfacing to their optimization algorithms such that a "layered AI" approach may be implemented where the a layer of custom

PLANNING COMPLEXITY – LEVERAGING THE POWER OF GENETIC ALGORITHMS

By Christoph Naber

When planning something – anything, there are always numerous factors, which need to be considered to derive a useful plan that is of practical value. This holds true for school timetables as well as for the build order of engines on an assembly line. A general overview of algorithmic solutions and approximations cannot be condensed down into one single book, let alone one article. This article, therefore, strives to provide a short and somewhat simplified example based on the aforementioned engine assembly line use case of Genetic Algorithms (GA) as a starting point for readers to explore their own use cases.

Considering the assembly line example in more detail. The following are relevant factors which need to be taken into consideration:

Generally speaking, final schedules need to consider these necessary restrictions to be a valid solution. However, there is usually much more to a valid plan. In the case engine production example, there are yet further, very specific restrictions to be considered. For example, some engines may only be produced in a particular time frame. Then there might be engines where it may be more efficient to build ten of the same type in sequence. Yet other engines might have the constraint that only three may be produced in direct succession due to their complexity.

It can be very complex to design and implement an algorithm that can automatically generate a perfect schedule for these types of restrictions, let alone configurable restrictions that can be switched on and off as required.

This is where GAs come in handy. With increasing data availability, demand for these types of approaches has increased dramatically over the past few years. In recent projects, we developed these types of algorithms with astonishing results.

In general, GAs try to make use of the idea of evolution as a problem-solving approach. Even though this approach is derived directly from the theory of evolution, it is possible to apply the principles to the implementation of GAs. By so doing we have implemented a framework that allows for a customizable execution of a GA in a very generic way.

In our implementation, the user must supply information pertaining to the optimization problem in two parts:

1. The implementation for a Genome. Its purpose is to hold all genes – the most simple building blocks – of one specific potential solution for the optimization problem. Think of this as a kind of "DNA", where all information about how to "build" a living being is held.

2. The second part is an EvolutionPackage that can perform four different tasks:

a) Assess the overall fitness of a Genome. Think of this as how well the Genome solves the task.

b) Select from the whole Generation those Genomes that have the most potential for further improvement mostly based on their fitness.

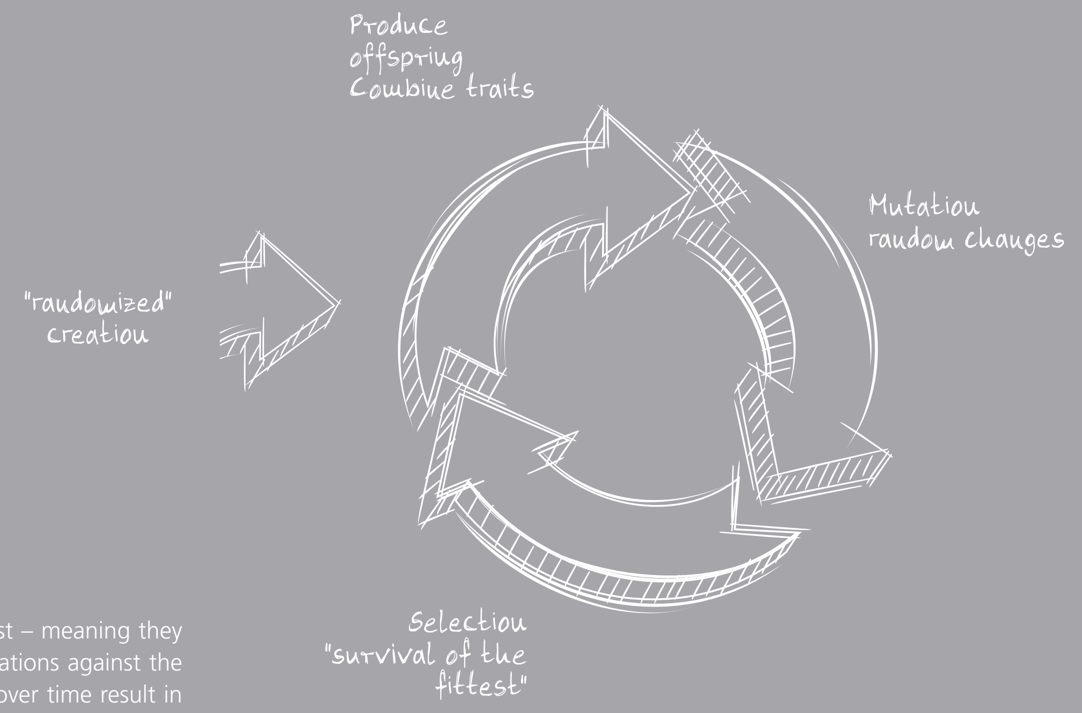
c) Randomly make changes on the Genomes to allow for improvement of these Genomes by chance.

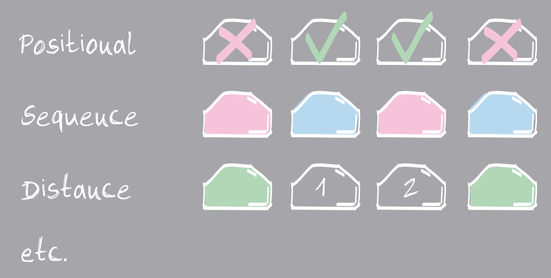
d) Combine the traits of two or more Genomes to a new one by selecting parts of the genes of one and parts of the other Genome. Think of this as the creation of offspring that inherits traits from its parents.

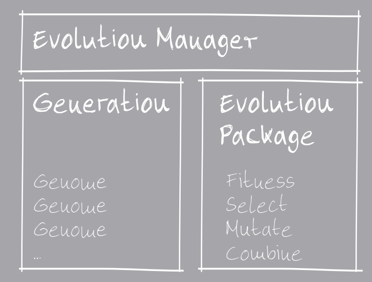
Admittedly, this is quite an abstract concept, so let us again apply this to our engine production example:

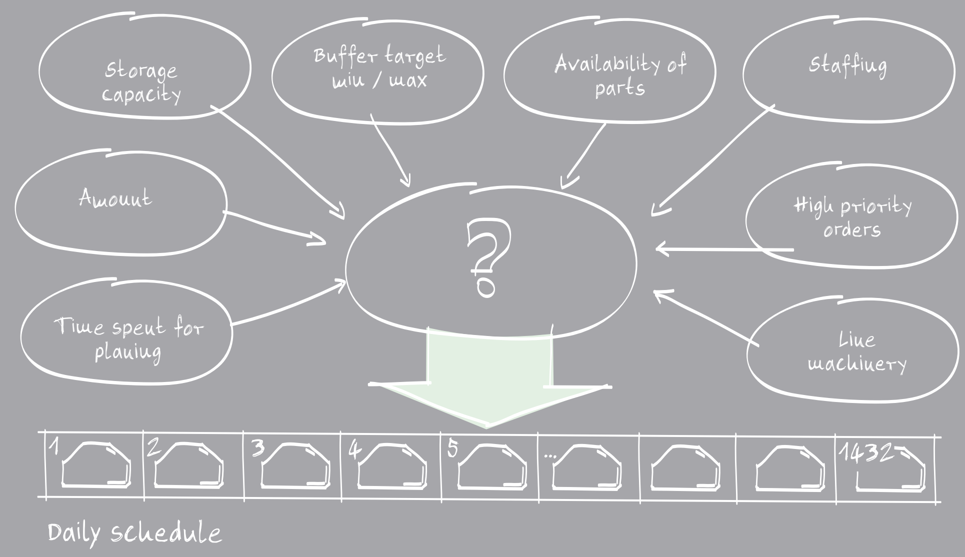
Representation of the engine production plan in a usable data structure was the first step as it should be in every project that strives to create a GA. Next, the genome is initialized with randomly generated sequences of engines based on production requirements. Due to the restrictions that had been defined upfront, the GA can determine the fitness of every randomly generated genome which represents a daily production schedule. These performance indicators are then used to select, randomly modify and combine the best of these schedules with an EvolutionPackage. This package addresses the unique demands of these schedules as they are permutations and therefore have to retain some properties even after mutation and combination phases. As the GA framework makes iteration after iteration, the selection of the schedules that perform the best – meaning they have the fewest violations against the set of restrictions – over time result in a final schedule that has no violations at all.

GAs are commonplace in the planning environment that large software-ecosystems offer. These pre-implemented algorithms allow for extensive customization but lack the total flexibility of our custom approach. Our self-implemented GA performed remarkably well in terms of runtime and quality of results generated relative to the implementation time. Our results show that, when looking for solutions to optimization problems, it is worth considering self-implemented solutions.









AI influences the operation of the optimization core. By so doing the inner workings of the AI becomes more transparent and maintainable.

In this layered setup exchange between the generic algorithm and AI-wrapper utilizes three methods:

1. Optimized Planning Parameterization
Weighting of planning criteria is a major influencing factor on optimization accuracy and performance. The AI-wrapper actively modifies and sets these planning criteria, thereby greatly optimizing the accuracy and performance of APS genetic algorithm.

Using the various data sources, e.g. manufacturing execution system and enterprise resource planning systems the AI-wrapper is trained (supervised training) to identify the optimal planning parameters for the APS optimizer under any given circumstances for which a planning run is to be carried out. The AI-wrapper will therefore use historic scenarios to produce a set of planning criteria in order to maximize the APS-Optimization performance and optimization result.

2. Production Sequence Abstraction
Optimization performance in all APS systems is heavily dependent on the complexity of activity networks (production schedules) and the optimization cost of individual activities within networks. Often complex networks are processed by the optimizer, in which the processing of individual activities is unnecessary and/or redundant. This has a detrimental effect on optimization performance. The AI-wrapper is trained to locate these redundancies and make suitable substitutions, thereby significantly reducing the optimization overhead.

3. Actual Activity Duration Utilization
Traditionally, optimizers use various forms of master data such as activity durations, in generating an optimized production schedule. However, especially in the case of activity durations, these sources of information are theoretical and deviations (positive od negative) in actual activity durations cannot be taken into consideration leading to inefficiencies. For example, if it were known that, certain activities can, under specific circumstances be performed faster; the optimizer could schedule more activities in the same amount of time.

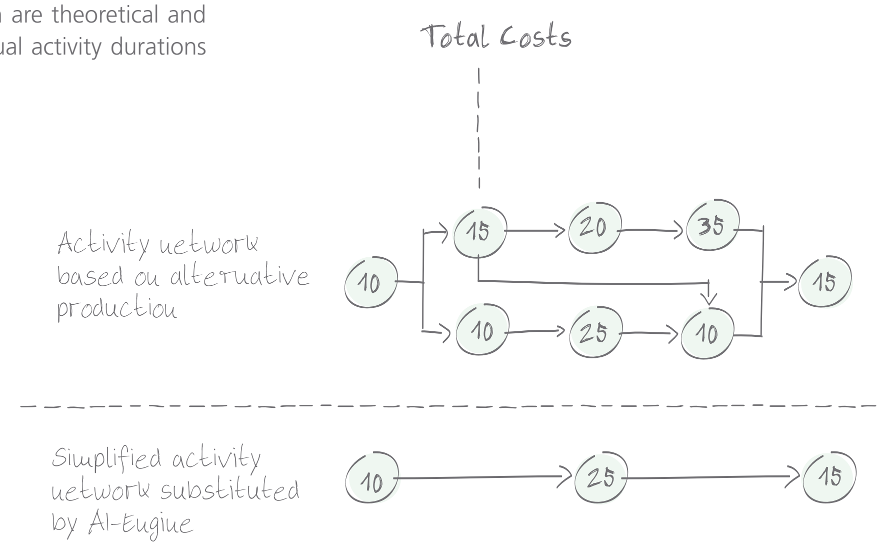
Actual Activity Duration Utilization does exactly this by allowing actual activity times as identified by the AI-wrapper, to be utilized by the APS-optimizer rather than theoretical activity durations as maintained in the master data.

Outlook: Closing the Loop

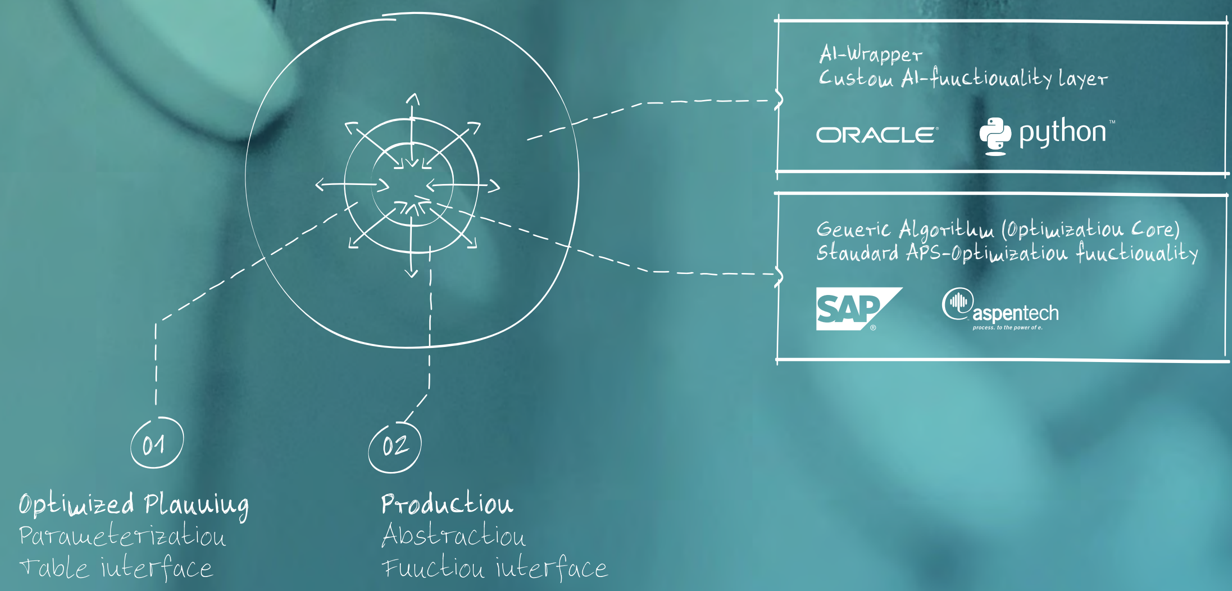
Artificial intelligence in production planning offers a variety of possibilities for advanced optimization and automation of planning processes. However, this is just the beginning. If one also considers the inbound side, that is the side before production, problems and challenges can be identified even earlier and solutions can be found much more effectively. For example, an additionally integrated Supplier Risk Management [4] can identify raw material and component bottlenecks including possible effects on production at an early stage and propose solutions by means of AI.

In addition, integration of the logistics outbound side in the "systemic loop" enables yet further potentials. For example, decisions can be made to solve production bottlenecks not only on the production side but also on the logistics side. Logistics and transport service providers can be integrated at an early stage to take into account restrictions and bottlenecks through an AI on the production side.

With complete integration of production, inbound and outbound logistics, artificial intelligence can be transformed from being merely an IT tool providing valuable functionality for individual business areas, to being the orchestrator of end-to-end value chain management with transparency and solution competence across a company's entire value chain network.



Reaching the next level. AI use cases are the key to reaching the next level of margin increase in the pharmaceutical industry.



Automating 1st Level Support

By Nikolaj Waller

IT-Self Service Augmented by AI

Today, corporate employees face many challenges during their work lives: complex tasks spanning multiple domains, an ever-changing environment, fast development cycles, to name just a few of these challenges. One typical but also daunting task is dealing with inevitable IT issues concerning hard- and/or software, such as password issues, VPN problems, or program crashes while trying to maintain high levels of actual productive work. When a disruptive IT-related issue occurs, employees usually have no choice but to call an internal helpdesk in search of assistance, only to be stuck in a queue for what may seem like an eternity. Having overcome this initial hurdle, employees are frequently forwarded within the support organization to service agents who, for cost reasons, are forced to perform several tasks in parallel. Helping employees deal with IT related issues – especially when these are non-standard problems, becomes an increasing burden, leading to discouragement and subsequent deterioration in the quality of service. This is especially the case for call center and helpdesk staff.

As is often the case, nowadays, automation and especially Artificial Intelligence (AI) can be used to alleviate these issues described, effectively helping all parties involved. Imagine the same scenario, this time with an automated 1st Level Support in place: The same troubled employee calls their company's helpdesk and is immediately forwarded to an agent to whom they can talk. The agent understands the problem, asks for relevant information, opens a ticket with the necessary information, and informs appropriate colleagues in the support organization. Now, you may say: "Great – finally the helpdesk is adequately staffed. What's the big deal?" Well, in the scenario described, the agent is an AI-driven chatbot, a custom designed and implemented software solution, providing key customer interaction elements that resolve several bottlenecks in our troubled employee's interaction with the helpdesk. Employees are no longer forced to waste productive time in queuing; they receive rapid problem resolutions and smooth transitions from focusing their attention on dealing with IT-problems to being able to continue their actual work. Helpdesk employees, on the other hand, do not have to face recurring repetitive IT tasks; they are free to dedicate their time and skill to complex, problems effectively, spending their time on more meaningful tasks. Overall, both the helpdesk and corporate employees encounter less friction and frustration in their day-to-day work thanks to AI.

On several occasions and in accordance with exact customer specifications, we have implemented such an intelligent, automated 1st level solution. The following sections aim to highlight not only the details and benefits of our approach but also provide insights into some of the challenges.

Human language is very complex, but it is also the most intuitive way for us to communicate, so given the recent technological advances in this area, namely the domain of Natural Language Processing (NLP), it seemed only natural to use a speech interface as the preferred way of communicating with our automation system. The challenges in implementing a speech interface arise primarily due to the way in which we humans communicate: we talk in a fluent, dynamic manner, extracting information from what was said, asking questions when necessary, and even switching subject's in rapid succession in-between. All of these factors pose immense challenges when implementing a chatbot capable of performing the aforementioned tasks.

Initially, we perform "context identification", a step in which information from what was said is extracted. Each company has its own terms, and these, in turn, have their own context. For example, "bank" can either be the financial institute ("I need to withdraw money from the bank"), or it can be the enclosing structure of a river ("We sat by the riverbank"). Some terms are even more ambiguous when it comes to a business context: "outlook", "lotus", "word". All these terms serve multiple purposes and could amongst many other things, be business tools you use daily. The technical solution to this is the so-called custom context identification using NLP methods such as word embeddings and keyword identification. By so doing, both meaning and context in which a term is used can reliably be identified.

Following information retrieval, we need to design the actual conversation. Think back to your last conversation. What did you try to achieve, and what did you say to achieve it? Did you get exactly what you wanted, or did you need to ask questions for further clarification? These are the exact questions we had to ask ourselves during the development of our AI models. Having found suitable technical solutions (details available on request) in answer to these questions, the result is a dialogue system, which feels natural. There is no pre-scripted path. Conversational twists and turns are taken by the system as necessary. Problem A has solution A, and thus path A is taken; problem B, on the other hand, results in a different conversational path, and our system can cater for this on the fly, in real time. What's more, our system can ask questions based on the information retrieved: if it does not have enough details for a specific case, it asks for more information. When an utterance is not understood, the system asks the caller to repeat themselves. This leads to a natural conversation similar to those many people are already conversant with from other settings such as Amazon's Alexa or Apple's Siri, resulting in high user acceptance.

However, there is more to the system than merely natural conversation since the ultimate goal is to provide a successful resolution to a service request, incident or problem. After having gathered enough information, the system carries out necessary actions according to the situation in hand, by creating a ticket in a ticketing (ITSM) system and filling in all available information in obligatory forms. Furthermore, our chatbot then classifies the tickets as service requests (e.g. resetting a password, granting user permission), incidents, or problems by either carrying out services directly or forwarding indents and/or problems, respectively on for further (human) processing.

The advantages of this approach are two-fold: whenever cases are encountered that can be resolved autonomously, the system directly carries out all predefined actions, thus reducing the helpdesk's overall workload. Incidents and problems that are more complex on the other hand are passed on to humans ensuring human involvement if and where it is necessary.

It is feasible that in the near future, more complex incidents and problems may also be broken down into steps that the chatbot and adjacent systems, such as RPA systems for example, can resolve -even without human interaction. Let's have a brief look into the future as to what further benefits such systems could then provide:

One significant advantage of chatbots is their continuous availability: Systems such as these can be active non-stop, 24/7, and can answer requests at any time of day or night. Since the application is developed as a service, it is scalable regarding number of users: it does not matter whether 1 or 100 employees are calling the helpdesk simultaneously since each and every request gets answered directly by the chatbot initially (for further details on scaling AI solutions, see also our A.I.dea Book article "Containerized Machine Learing Architectures"). Furthermore, with the advance of machine translation we can tackle problems faced by international, multi-language corporate environments. No matter how many languages employees speak, the chatbot will be able to understand them all since this system can be written language-agnostic or be implemented, including a translation component to cater for other languages. This allows a true single point of contact for any IT problem across multiple countries or simply supporting different languages and dialects in a single country.

Let us briefly recap what we have been discussing: a chatbot in combination with ITSM (Ticket) Tools such as has been implemented by MHP, is an excellent opportunity for automating first-level support since non-complex and repetitive tasks can be handled by AI to alleviate stress on human agents. To do this successfully, it is necessary to create a flow in conversation, which is as natural as possible, allowing for a system that is both user-friendly and scalable in both number of parallel user interactions as well as the number of languages supported. Due to the usage of natural language in our dialogue system, an intuitive way for users to interact with the system is provided resulting in an immense adaptation potential of such systems. Because the advantages of these types of systems are so immediately apparent, we at MHP predict widespread usage of this type of technology within the next five years.

Customers want new, digital alternatives to telephone services. The added value of chatbots is greatest in simple and clear processes, when responses are personalized and show signs of empathy.

William Cobbah | Artificial Intelligence Business Developer

Strategic Recruitment of Applicants

By Johannes Stark

Identifying High Potentials

The greatest challenge in the recruitment of human resources is to find competent people with high potential who are best suited to meet a company's requirements. To achieve this, recruiters essentially have two main talent pools at their disposal:

- Their own employees in the context of internal personnel recruitment
- External potential employees

In terms of initiating a dialog with potential external candidates, this can either occur by unsolicited or specific position targeted applications directly from candidates or as a response to a company proactively approaching a candidate (headhunting). Especially with applicants who are acting on their own accord, it is important to identify high potentials rapidly to be able to provide feedback within a short space of time, reducing the probability that these valuable candidates will look elsewhere. These time constraints are a heavy burden on the daily routines of recruiters around the globe, made worse by a growing volume of applications due to an ever-increasing number of recruitment-platforms, -portals, and social media acting as input channels. The high number of applications, including CVs and application letters that have to be screened by recruiters every week means that recruiters spend much of their time pre-selecting and initially evaluating candidates. As a result, they often lack time to conduct timely job interviews to get a personal impression of the applicant. This, in turn, leads to a lower probability of success with high potentials, since these are often recognized too late or not at all. Ensuring consistent quality in the evaluation of applicants is yet another challenge for recruiters. Depending on the mood and mental state of a recruiter, a large number of applications will inevitably lead to subjective fluctuations in the evaluations. In addition to this, recruitment departments in most countries are obliged by law to adhere to regulations regarding equality or equal opportunities. In Germany, for example, this is the General Act on Equal Treatment (AGG).

Most notably, enterprises want more efficiency and higher quality in the selection of applicants. Above all, the time span between receiving the application and the provision of feedback is crucial, since it is a well-known fact that companies with short feedback periods have an increased probability of hiring high potentials and/or candidates of choice. The bottom line is; optimization in the recruitment process can only be achieved by focusing recruitment resources on high potentials and candidates of choice.

Automation with Text Mining and NLP

The primary aim is to automatically compare the required qualifications, competence and experience for an advertised job with the characteristics, skills, and abilities of the applicants to identify the most suitable match. The information received from applicants is contained in unstructured form as text data from CVs and application letters. In an initial step, this unstructured text data is extracted from the CVs and application letters partially by the use of optical character recognition (OCR), which is enriched with structured data acquired throughout the application process. Natural language processing (NLP) methods like tokenization, stemming, and stop word filtering are used to prepare the data for analysis. Based on the job profiles in question, the required qualifications, competences and experiences are compared with the processed text data. Clearly, for these purposes, NLP methods employed must be able to take spelling, foreign languages, synonyms and associations into account. This can be achieved efficiently by applying MHP's adaptive NLP Pipeline for example (see the article MHP adaptive NLP Pipeline for further details).

The results of NLP are used with the aid of algorithms based on similarity measures. These algorithms identify candidates who are most suitable for a given position, guaranteeing an objective determination of high potentials, since numerous soft and hard facts are derived for each applicant purely on a data-driven basis. Core to these matching algorithms are similarity values which, for example, are derived by word embedding from the applicant data and job requirements. Using job-specific weights, the similarity values of soft and hard facts are consolidated into an overall assessment of the match between applicant and job.

The objectivity of the algorithms is ensured by the fact that they neutrally evaluate the data of all applicants and do not take into account information such as ethnic origin, gender or age. The unstructured text data, as well as the structured applicant data, are recorded directly in the application system, processed in the described manner, with the results being automatically written back into the system.

Increasing the Probability of Success and Saving Time Simultaneously

By implementing the automated analytical approach for identifying high potentials in an existing recruitment process, the effort of pre-selection and initial evaluation can be reduced by up to 90%, based on experience in several relevant MHP projects. This frees up recruiting resources, which can be used, for example, to schedule, prepare, and conduct interviews.

The average feedback period for applicants, especially for high potentials, could be reduced by up to 80% in previous relevant MHP projects. This significantly increases the probability of successfully recruiting the most suitable applicants with the highest quality and potential. The objectivity of the algorithm enables a consistent recruitment process and ensures an assessment of the applicants that is free from prejudice if systematic biases can be excluded (For more on bias in this context, see also the section titled "The Human Factor" in the A.I.dea Book article "First Do No Harm: Why AI Ethics Matter").

The areas of application for such algorithms are enormous. On the one hand, they can determine which high potentials are available for a position among the applicants for the post in question, but since work-effort is no longer a limiting factor, applicants who have actually applied for another job can also be taken into consideration. On the other hand, the system can also be used to run checks to identify which of the currently vacant positions an applicant is best suited for, regardless of which position the candidate has actually applied for. Adaptation of this approach to include internal or proactive recruitment does not pose a problem.

> **More than 80% of the prevailing data is not structured. Sooner rather than later, it will become imperative for business front-runners to use this potential in their business decisions – and it's a great time to start analyzing text data: Recent advances in deep learning now make knowledge extraction from fuzzy data such as texts much more feasible.**
>
> **Vanessa Viellieber**, Senior Data Scientist & Natural Language Processing Expert

MHP has helped several customers by increasing the efficiency and quality of their recruitment process through the use of NLP methods and matching algorithms. MHP Machine Learning Engineers, Data Scientists, and Software Developers advise and support their customer's end-to-end from conception, through development and validation of algorithms, to the implementation and visualization, adding value through increased efficiency to existing recruitment processes.

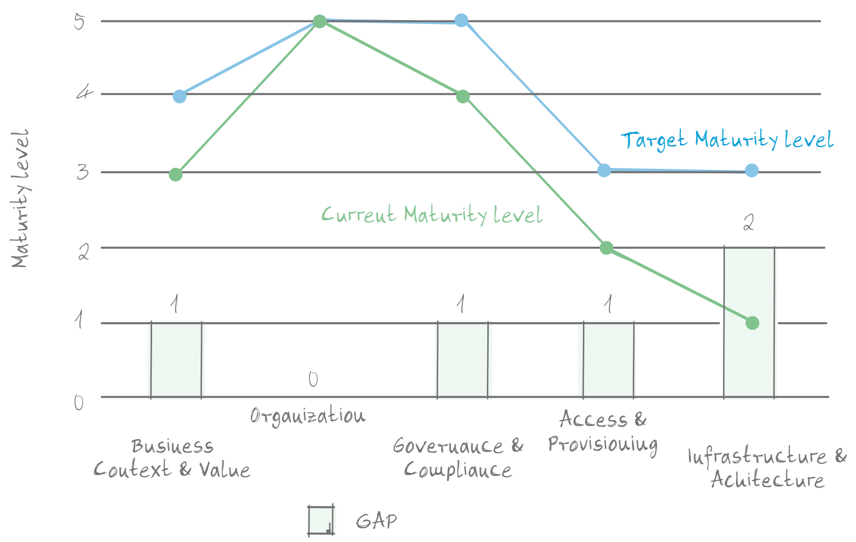
METHO-DOLOGY

03

Data Maturity Assessment

By Michael Genne

Figure 1 – Measures can be derived from the deviation between the current and target maturity levels

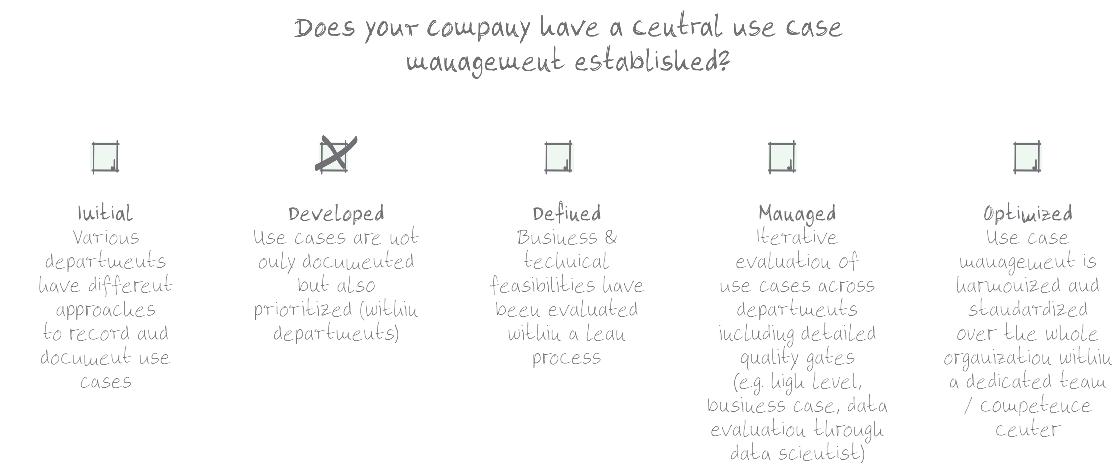


Nowadays, companies are aware that data has become a crucial part of any enterprise's value chain, potentially providing an important strategic competitive advantage. Processing raw data and transforming it into valuable information by utilizing Artificial Intelligence (AI) can improve efficiency and reduce costs while often generating additional revenue.

Nevertheless, Big Data has proved to be a disappointment due to the high failure rate of Big Data projects which remains unchanged to this day [1]. In our experience, the cause of failure in Big Data projects is not the availability of technologies, but rather the lack of integration of existing business processes and systems, internal politics, lack of skills as well as various governance challenges. To assist companies in facing these challenges, MHP has developed an interdisciplinary approach to accompany its customers on the journey along the data and AI value chain, from the definition of a strategy to highly scalable and productive artificial intelligence solutions. This Data Strategy Framework is described in our Article "How to successfully accomplish your Data & AI journey". By employing this approach, the usual obstacles encountered during the transformation towards a data-driven company can be recognized and therefore avoided, which in turn minimizes risks during transformation. But what exactly do we mean by the term "data-driven company"? More than just installing the right tools and applications, data-driven organizations treat and manage data as a primary asset just like human resources or financial assets [2]. In this role, data supports business processes and decision-making as well as the enablement of digital business models.

To become a data-driven company, organizations need to understand their status quo in terms of data maturity, and whether or not, they still have "data maturity white spots". An increase in data maturity is the evolution of an organization towards being able to integrate, manage and leverage all relevant data sources. It is not merely about having some technological components in place to deal with unstructured or high volumes of data. It is about the ability to make those high volumes of data, be they structured or unstructured, answer business questions and thereby creating value. Entire business cases are created based on data providing cost reduction

Figure 2 – Our web-based survey enables the evaluation of strengths and unused potentials with regard to Data Maturity.



or enabling revenue targets to be met. In other words, data maturity is a journey that involves creating an ecosystem that includes data management, analytics, governance processes, technologies and organizational components with the aim of generating business value.

At MHP, we have created a Data Maturity Assessment in response to very specific requests from various organizations to understand their current data maturity in the transformation process of becoming a data-driven company. The Data Maturity Assessment includes a web-based survey that guides users through areas (e.g. "business context & value" or "governance & compliance") and their associated fields of action with particular questions.

Our maturity assessment provides guidance for companies that are at the beginning of their Data and AI journey since it gives a structure for a data program, highlighting the most important areas of action in the context of data driven transformations and determines where exactly to start. Moreover, the maturity model offers a methodology to measure and monitor the state of projects and programs, giving clear advice regarding the necessary steps and associated work-effort required to move to the next level of data maturity.

The objective of the maturity assessment is not only to determine the position of a company or department on the Data & AI journey, but also to support business and IT leaders to understand and evaluate the essential enabler areas for accelerating a company's data transformation process. Although the purpose of the assessment is to assess the data maturity level in the new, enabler areas (detailed below), we are convinced that it is just as important to consider and leverage existing capabilities and to invest in fields of action that best accelerate the progress along the data journey. We achieve this by helping organizations understand best practices used by companies that have already successfully undergone the data transformation process.

MHP's Data Maturity Assessment consists of five maturity levels: initial, developing, defined, managed, and optimizing. As organizations move through these levels, they gain more value from their investments. The assessment measures the maturity of a data program across five enabler areas and corresponding fields of action that are key to deriving value from data:

- Business context & value
- Architecture & infrastructure
- Data governance & compliance
- Data access & provisioning
- Organization

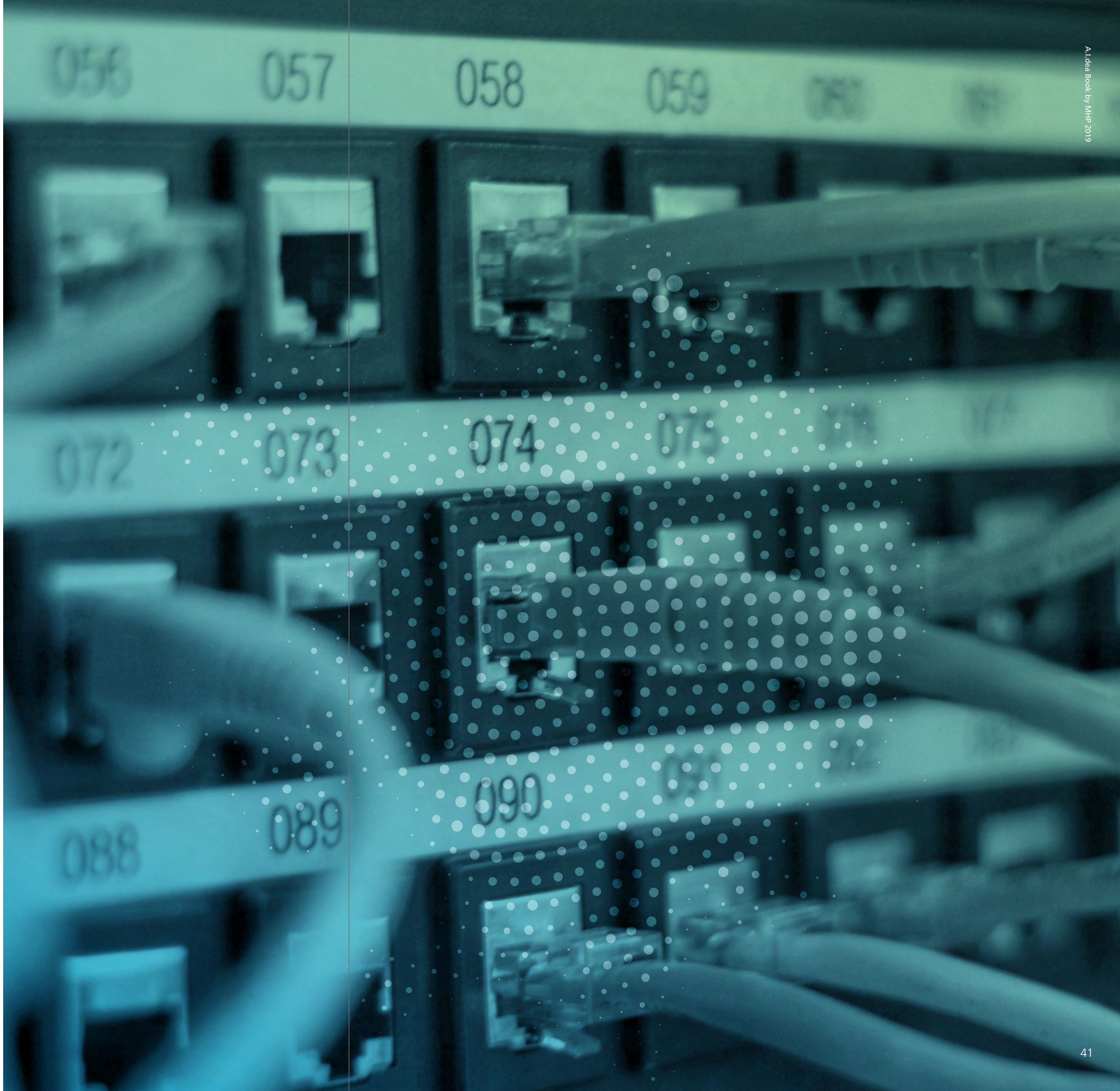
To facilitate progress on customer Data and AI journeys, we analyze these five different enabler areas holistically across the organization, then examine the corresponding fields of action per enabler. Finally, we measure the maturity level for each field of action to understand the overall data maturity of an organization across all areas. Fields of action in the enabler area "business context & value" are, for example, a well-defined data strategy (i.e. one that adds to, and is in line with, the company's overall business strategy), use case ideation processes as well as a standardized use case management. Standards, processes and KPIs for data quality and master data management are fields of action in the area of "governance & compliance" and standardized data models or a scalable infrastructure are examples for "architecture & infrastructure".

Based on the survey results, the deviation between current and defined target maturity level for data programs are identified. MHP derives measures to reach the next level of maturity. This can include the implementation of rapid prototyping approaches to verify identified use cases as an example of a measure from the enabler area "business context & value". Data governance roles and responsibilities such as data stewards and data custodians oversee the life cycle of data assets. The definition of such roles as well as of data management principles increases the maturity in the enabler area of "governance & compliance". In the "organization" area, one example is the development of a committee structure that focuses on data strategy and projects for the strategic and operational control of a company-wide data platform. The development of an authorization concept for a data platform serves to regulate data access and to ensure the "need-to-know" principle, which leads to an increase in maturity within the enabler area "access & provisioning". These measures are combined in a structured roadmap with iterative development stages since data transformation is a change process that can last several months or even years.

In our experience, there are two approaches to becoming a data-driven company, the "sustainable" and the "acceleration" path. The sustainable path starts with the identification of data use cases, which are translated into functional and non-functional requirements in order to derive the target architecture of a data platform as sustainable foundation of current and future Data and AI use cases. This path considers organizational questions like collaboration models or communication concepts as well as data governance matters like data quality or meta data management. The acceleration path is an alternative entry point into the AI value chain, where use cases are implemented in a lean and scalable manner in a short period of time using modular automation approaches (e.g. continuous integration and continuous deployment). However, the two approaches are not in competition with each other but complement each other depending on the objective pursued (speed and technological scalability versus a holistic foundation).

As mentioned at the beginning of this article, Big Data has been an enormous disappointment, and the data-driven transformation raises many questions and uncertainties. But it also offers countless new opportunities. Reaching maturity and becoming a data-driven organization requires many iterative development stages, organizational change, and a new mindset.

We continue to accompany organizations on the journey to seize these opportunities ensuring that data-driven transformations are successful.



How to Sucessfully Embark on Your Data & AI Journey

By Dominik Graetz

Experiences and Key Learnings to Emerge as Front Runners in the Digital Transformation Race

Everyone wants to use the fascinating, promising chances and opportunities associated with one of the most controversially discussed topics within the last few years: Artificial Intelligence (AI). AI may enable potentials like identifying and optimizing cost-saving potentials, improving quality issues or increasing the customer value based on personalized products or services. According to a survey of more than 2.500 executives, seven out of ten companies perceived little or no benefits from their AI initiatives. It seems that despite the massive potentials, many initiatives are still not delivering the expected gain. 90% of the interviewed executives, however, agreed that AI represents a business opportunity, but 45% also fear the risks associated with AI [1].

The concept of Artificial Intelligence, however, is not as easily accessible as some more traditional IT-concepts; therefore the question is: Where does one start?

- Break down corporate strategies and formulate a data & AI strategy?
- Implement initial POCs to get a feeling as to how a process can be improved using AI?
- Implement a data platform as the most important foundation for AI applications?

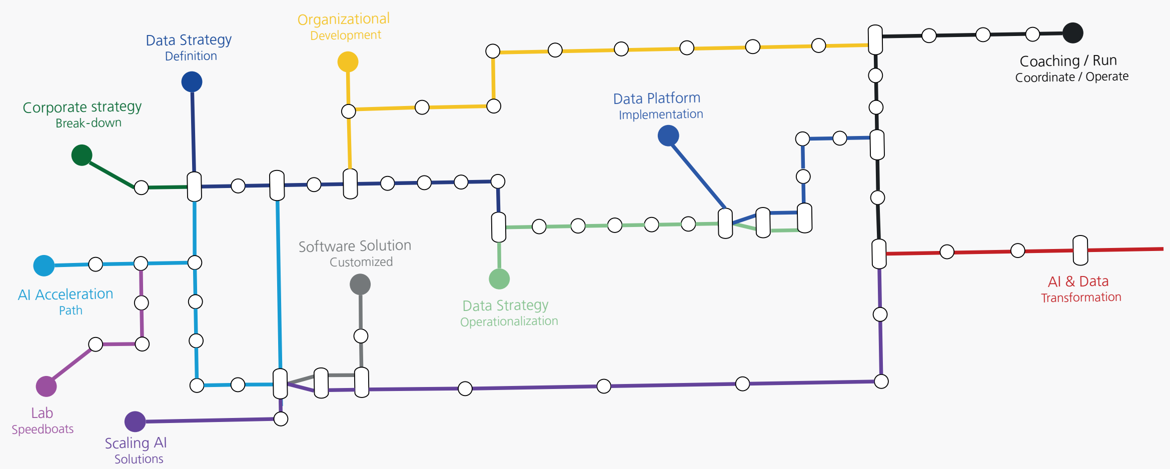
All these questions pose valid starting points, but one should bear in mind that becoming a data and AI-driven company is not just a one-time effort, it is part of a transformational journey, which cannot be traversed overnight.

Imagine you are in London getting into the tube to reach your desired destination, which in our context, is the AI and data-driven transformation (figure 1). There are multiple routes to get from a given starting point to any destination. Some may for example require tackling disciplines like strategy, organizational changes or scaling AI software solutions. All of them however are ultimately stops and changes on the journey that is the AI and data driven transformation.

As each company has different needs, it is important to evaluate the data maturity of the current organization. Data maturity can be described as the evolution of an organization to integrate, manage and leverage all relevant data assets. Key components which should be evaluated are for example: Main business value drivers (e.g. use cases, process optimizations), architectural and infrastructural foundation (e.g. data platforms, data lakes), the necessary data governance aspects (e.g. data governance roles, data quality processes) and the development of collaboration models (e.g. agile, scrum) as well as data-driven processes within organizations. Depending on a company's data maturity level and current strategic goals, various entry points on the Data and AI journey are feasible. Those challenges are then translated into bespoke recommended actions, based upon applied practices (See also our A.I.dea Book Article "Data Maturity Assessment"). To provide insight into the types of challenges companies experience during their journeys, consider the following scenarios:

A.I.dea Book by MHP 2019

Figure 1 – Data and AI journey as your timetable for the AI & data driven transformation



Imagine one is shaping the digitization of a medium-sized manufacturing company delivering high-quality products. The interim goal is to generate data from existing production and enterprise resources, like machines, sensors, and ERP systems. But generating data alone does not create value; therefore one should rather continuously exchange information to improve learning processes and sustainably increase the overall performance, closing the loop between development and manufacturing [2]. Within the example of a medium-sized manufacturing company, one has to deliver tangible solutions in a lean manner under immense time pressure, to prove, that AI is helping to support a company's strategic targets. Our experience shows that projects such as these are challenging as expectations for AI solutions are always high. Despite this, investment capabilities remain somewhat limited. To resolve this contradiction, one has to start from a business perspective by challenging existing use cases or identifying new potentials. Prioritization is essential since budget limitations, insufficient infrastructure, technological components and lack of experienced workforce mean that the proverbial "first shot" needs be on target. Fortunately existing methods and architecture, like CRISP-AI [3] and AI services architecture are available and can be utilized based on the defined use case and requirements to develop an AI POC infrastructure. This type of platform is of manageable complexity, while it is designed based on flexible and integrable components like Docker and Kubernetes to enable scalability within a corporate environment (see also our MHP A.I.dea Book article "Containerited Machine Learning Architectures"). These technologies are also accessible through major Cloud providers like AWS and MS Azure and have persuasive advantages, like pay-on-demand, scalability, and accessibility - ideal for scenarios with limited resources pressured to deliver results within a short period.

Figure 2 – AI Acceleration path to deliver prioritized use cases fast and lean

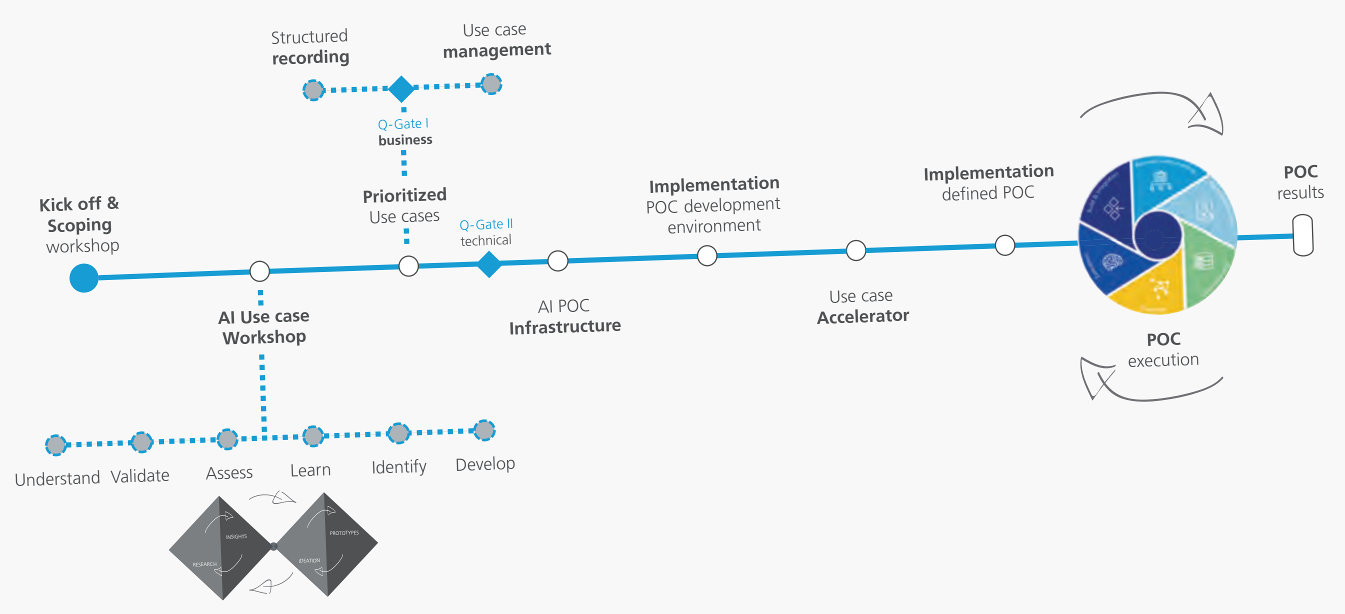


Figure 3 – Data strategy framework as foundation for a data platform, and its correlating organizational and governance requirements

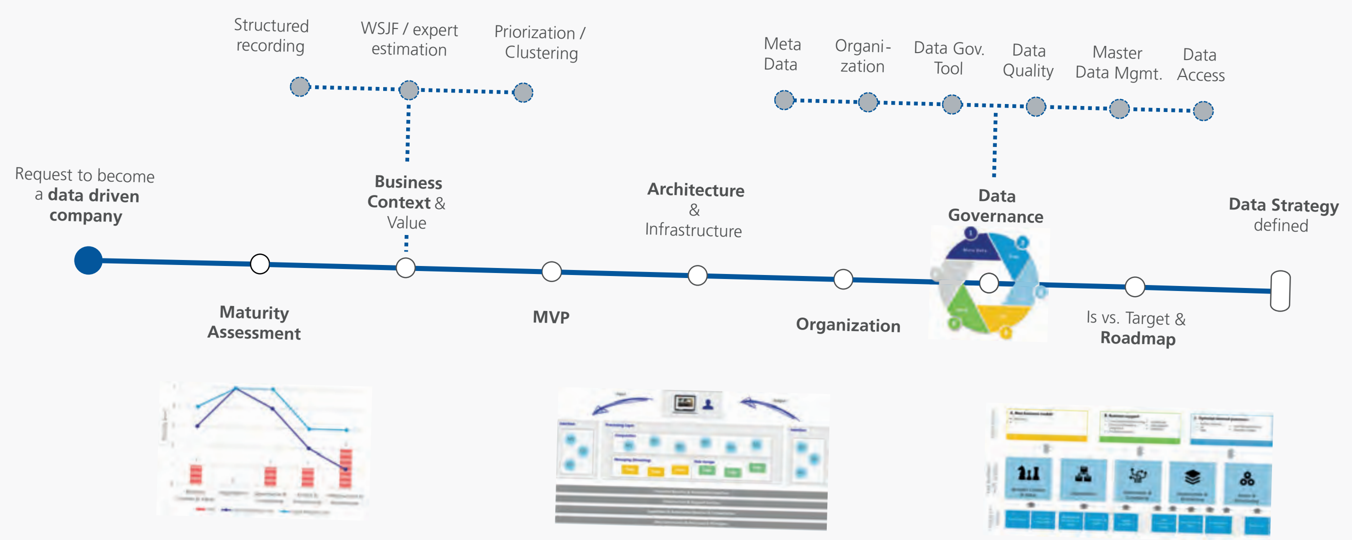


Figure 2 depicts a so-called "acceleration path" which includes various stops along one's journey, geared towards rapid deployment of AI solutions. It includes AI use case workshops as an ideation phase, structured use case management to continuously fill and prioritize use case pipelines, as well as the implementation of a lean POC development environment in which to implement the defined use cases. These stops are crucial on route to successfully running POCs, which can easily be transformed into scalable minimum viable products (MVPs) and ultimately into an enterprise software solution (Figure 2).

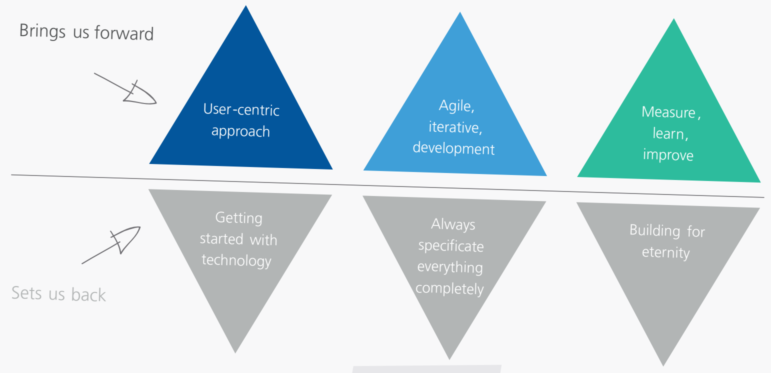
Imagine working in a B2B bank where data is distributed among different systems and more than 15 use cases, like finance forecasting provisions or consolidated management reporting using, e.g. Python. Direct implementation is not feasible due to the decentralized distribution of data in data silos. In this case a harmonized, company-wide data strategy needs to be established, providing guidelines, breaking up the data silos, and supporting interaction between various stakeholders (business and IT) to harmonize data assets and discuss core concepts of the data & AI-driven transformation. In this scenario, the primary goal is not to development scalable AI services but rather to develop a data strategy and implement a data platform, which is the foundation for analytics use cases and future AI services.

The starting point is to record, evaluate, and cluster use cases to derive functional requirements for the data platform. These use cases are the foundation for an MVP always bearing in mind that the solution's architecture and infrastructure must be flexible enough to scale into a fully operable data platform. Various technological components can be used for the data platform (technology-agnostic approach). Examples are SAP HANA if we are using only structured data, or Hadoop components in case the requirements include unstructured data. Once the technical questions have been answered, organizational issues regarding collaboration models, agile approaches, and knowledge composition within an organization should be approached. For the target set-up, processes, data governance methodologies, and frameworks need to be established or modified to enable employees to work efficiently without violating policies or regulations. In banks one key challenge lies in fulfilling regulatory requirements while maintaining access, permissions and change processes flexible enough not to drown data analysts and –scientist with administrative overhead. Some banks have even taken this one step further and have successfully adopted DevOps in core banking processes, so they were able to scale from a few hundred code deployments per month in 2012 to over 18,000 code deployments per month in 2014 [4]. Journeys like this show how important it is to start a transformational journey on a strong foundation. Utilizing a step-by-step approach to get first results quickly while iteratively testing and adapting results in a user-centric solution offering real benefits.

Both of those examples are only two possible legs on a journey towards the AI and data-driven transformation (Figure 1). Another route may include tackling the transformation by breaking down the corporate strategy and building the foundation for the establishment of a data strategy. Other possibilities would be re-structuring the organization or parts thereof, towards becoming an agile and DevOps driven organization. These values are crucial to be able to uphold the massive acceleration potential of such a transformation.

Every organization has its individual challenges, defining the starting point on the Data and AI journey. Nevertheless, all organizations have in common that they need to solve business challenges. As discussed in the MHP A.I.dea Book article "Panacea of AI: Advantages and Limitations of Deep NNs", AI (or, more specifically, Deep Neural Networks) is not the universal solution to all modern-day business problems. That is to say; it is vital to keep an open mind especially with regards to other technologies, approaches, and concepts.

Figure 4 – Data thinking principles to generate Data & AI use cases



Several partners and customers have reached out to us with their backs against the "AI-project-wall" such that we were able to observe that amongst other causes, the top reasons why AI projects fail are:

1. Putting Technology First

Purchasing a high-gloss out of the box "AI platform", which at first sight is a perfect match for a company's requirements and can be up and running in a matter of days, is often deemed irresistible. The well-established principle that "out-of-the-box" is seldomly a good enough fit to be able to add real value is true in this instance. For many years pre-build software solutions were customized to specific customer needs. Nowadays, due to ever-shortening development cycles and increasing requirement complexity, such an approach is seldom agile enough for modern software development. We, therefore, recommend a technology-agnostic approach that identifies functional requirements (e.g. real-time data transfer, messaging services or visualization in tailored frontends). Customer-specific use cases lead to tailored AI solutions, which can be generated in individual AI use case workshops — utilizing creative methods like data thinking to get user-centric solutions that solve real-world problems. For prioritized use cases, relevant technologies are identified based on criteria like cost-saving potential, potential for quality improvement, implementation complexity and job size. This step by step approach ensures early returns on investments while maintaining the overall direction according to continuous use case management, which is aligned with the campany's overall strategic goals.

2. POC "Madness"

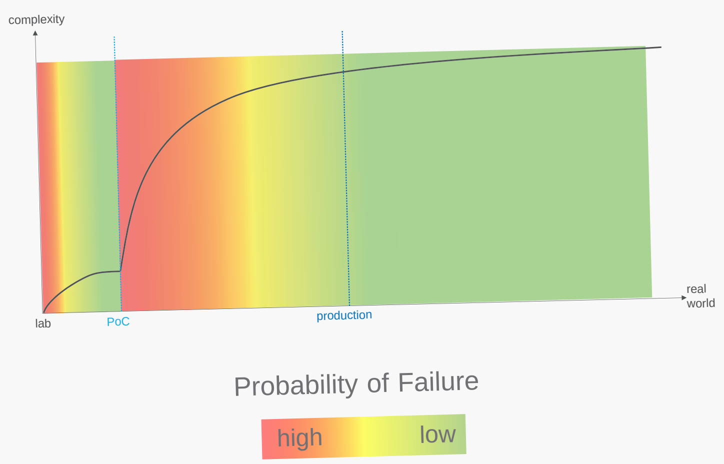
As mentioned earlier, one first step could be to implement initial POCs to rapidly develop an idea as to how a process can be improved using AI – if at all. However, POCs should be created to address a specific and apparent business need and not just for the sake of it! [5] We have found a lack of purpose to be a significant factor in the failure of AI projects. Also, more than 80% of all POCs will never see the productive stage [6] and therefore deliver no tangible value for companies. During a POC, synthetic data is generated to prove a point, i.e. to get successful results, but real-world data is often very different.

Also, very often, the architecture of AI POCs is not sufficiently well designed to allow scaling into a productive environment. This then begs the question of why so many companies are still following the POC "madness" path?

Based on the experiences of the last few years, the reasons are manifold. Sometimes an organization itself is driven by the fear of AI, as employees are afraid that AI may take over their jobs. However, as of today there is little or no evidence to support this. By contrast it is suggested, that humans and machines have distinct strengths and weaknesses, when predicting certain outcomes [7] and it is rather this factor which will determine the distribution of future jobs. Sometimes there is also a lack of management commitment towards AI solutions. As mentioned before, the concept of AI is not as tangible as other IT concepts and enabling traditional IT departments will require not only monetary investments but also an investment into new skills and a change in culture. Christopher Little, a software executive, recently said: "Every company is a technology company, regardless of what business they think they're in. A bank is just an IT company with a banking license". [8]

In other instances companies do not realize that the focus and majority of effort in implementing integrated AI solutions is software engineering and not data science. Software engineering requires different skills and experiences, especially when a POC needs to be integrated into an existing corporate environment. True value can only be created by an integrated software solution and not by a successful proof of concept.

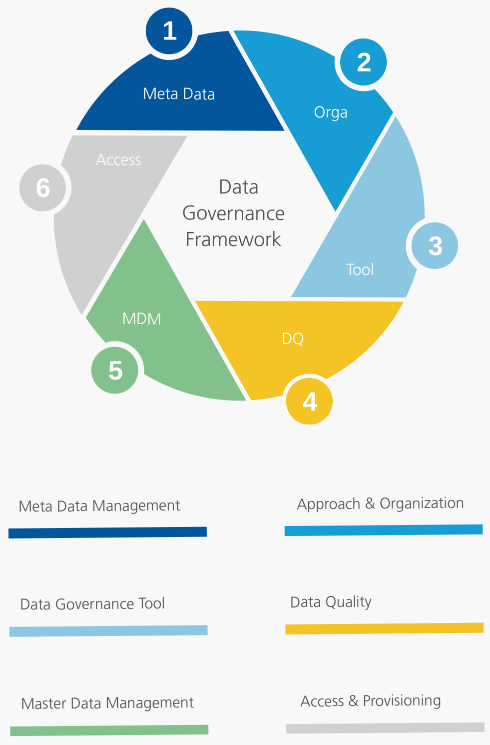
Figure 5 – Probability of failure of AI POCs



3. Insufficient Data Governance

Many companies want to use data to improve business processes and enable new digital business models and strategies, but do not treat data as a primary asset, such as HR, production equipment or other capital assets. Data management and data quality have been identified by recent studies as the main problem areas when introducing AI / advanced analytics and business intelligence initiatives [9]. Data governance is still regarded as being a source of cost and complexity. However, this view needs to be adjusted in order to move forward in business transformation. Hardly anyone would argue with the fact that it is crucial to invest in human capital like employees. Quoting Peter Baeklund's famous statement regarding investing in people: The CFO asks the CEO: "What happens if we invest in developing our people and they leave?" The CEO replies: "What happens if we don't, and they stay?"

Figure 6 – Core disciplines of data governance



Our data governance framework describes key disciplines, which need to be tackled to treat data governance issues within an organization (Figure 6). Let's take a brief look at three of the six key areas where we identified the most significant issues in recent years:

Metadata management, which includes the definition, specification, and information on key attributes of data. For example, for a customer, an extract of the metadata may include:

- description / definition of a customer
- data type (e.g., text)
- data format (e.g., VARCHAR 32)

Useful metadata is particularly relevant as data analysts and scientists have to work with data from various data sources and need to understand the data and metadata model to generate value with AI solutions.

Data quality is one of the most crucial factors if you want to fuel an AI solution with data, as the quality of data strongly correlates with the predictive quality of the model. Improving data quality is not a one-time effort as the amount of data is always increasing, so the quality assurance cycle of identifying, acting and measuring needs to be continuously executed. If data is not available in the required quality, then at least for POCs it needs to be synthesized, fueling the effect previously described as POC "madness".

Access & provisioning need to be regulated by comprehensive authorization concepts, which adhere to "need-to-know" principles or regulatory and GDPR requirements. Previous experiences have shown that designing suitable authorization concepts for the types of applications in question can be quiet challenging. As different systems with different data owners have their own functional and non-functional requirements. Furthermore, should a provisioning concept be based on data classification and the depth of access is based upon roles. A data scientist will need to be provided with more liberal access rights than a business analyst, as he/she needs to work with more extensive and "rawer" data sets.

The mentioned data governance dimensions all have a common denominator, as all of them need to be managed by business and IT. They can be supported by technologies, like data governance or data quality tool, but the tools themselves still need skilled staff like data stewards (business) or data custodians (IT) to gather, manage and maintain its content.

Bottom Line

Companies should carefully consider some of the learnings mentioned above while undergoing their personal Data and AI journeys since these can be decisive factors in emerging as front runners in the digital transformation race. Lots of AI solutions are still focused on enhancing existing processes in various areas, like manufacturing, marketing, pricing, etc. In some cases these improvements are equivalent to improving the fuel consumption of an internal combustion engine in an age of electrification where new mobility concepts and autonomous driving cars are emerging. Visionaries and executives should, therefore, focus on the reinvention of rusty processes and structures, as whole businesses and industries are disrupted along the entire value chain. Only then can the true potential of AI be leveraged: Not in blindly adapting existing processes, but in developing new things at large [10].

User-Centered AI

By Thillai Sivakumaran, Katarina Preikschat, Sebastian Große

Merging Design and Data Science to Generate Innovative Solutions Users Actually Like.

Why Do AI Projects Fail?

It all seemed very promising when IBM announced that its IBM Watson AI Health could support cancer treatment using AI [1]. It was to be the revolution in medicine, but the project was stopped by the MD Anderson Cancer Center of the University of Texas in Houston after $62 million was spent on the development of the so-called "Oncology Expert Advisor": The results obtained by the AI were "useless and sometimes [even] dangerous" for the treatment of cancer [2].

The recommendation system for customer service representatives at Mr. Cooper, the largest non-bank mortgage provider in the U.S., was also classed as one of the most significant failures in AI project history. The company expected substantial cost savings from proposed solutions to customer problems, but after nine months of development, customer consultants decided not to deploy. The reason was that AI did not offer any suitable solutions. [3]

These are only two of a much larger number of failed AI projects. Pactera Technologies, NA Inc, a global technology company, and Nimdzi Insights, a consulting firm, found that 85% of all AI projects "ultimately fail to deliver what they promise companies" [4]. It is often not the developed AI that fails, but the most overvalued or unrealistic expectations of project members and decision-makers that lead to assessed failures. These expectations may have been based on insufficient AI expertise and too extreme a focus on sales and KPI targets. In the example provided at the beginning of this article, IBM Watson AI Health was trained with hypothetical and artificial patient data so that its decisions were not based on a wide range of real cancer cases, but on the treatment preferences of a few physicians [5]. Therefore, the Oncology Expert Advisor is not applicable in a real-world setting. Similarly, it was not Mr. Cooper's algorithm that failed, but the data with which it was trained: Instead of using real problem descriptions and customer formulations, detailed technical data that customers would not use in their own language was used [3]. The recommendation system was therefore unable to provide appropriate solutions to real customer problems.

To prevent AI projects from failing after months of effort, we have developed a best practice process. Our user-centric, problem- and data-oriented approach helps find real business problems, focus on realistic goals in order to meet expectations. With the "fail, before you start" approach, we turn projects into successful projects.

What Value Does Design Add?

We believe in user-centric solutions. But what exactly do we mean by this? In the true spirit of Design-Thinking, our approach starts by investigating our target group's problems and challenges. In essence, we want to develop empathy for them, understand them and identify specific issues. We derive particular problems from previously formulated rough problems on the basis of our research.

Specific problems pose challenges to be resolved in due course. To do this, we leave the "problem-space" and begin with the development of new solutions in so-called "solution-space". At first there are no restrictions, which limit our solution-space. On the contrary, we develop as many ideas as possible, which we can then compare and evaluate.

Once we have prioritized individual solutions, we need to make them tangible for our potential target groups and stakeholders as quickly as possible in the form of simple prototypes. A simple prototype is sufficient to illustrate the prioritized solution and further develop it with our target groups. The underlying aim remains the validation of assumptions at an early stage, iteratively optimizing the solution while meeting the requirements for rapid implementation. We often encounter the challenge of technologies raising expectations that ultimately cannot always be met. Furthermore, new and sophisticated technologies such as AI are usually not included in the solution-space due to a lack of user expertise and design thinkers. These challenges must be overcome by incorporating design and data science right from the start.

How Do Design and AI Work Together?

"Despite its name, there is nothing 'artificial' about this technology – it is made by humans, intended to behave like humans, and affects humans. So if we want it to play a decisive role in tomorrow's world, it must be guided by human concerns" [6]. This remarkable sentence by Fei Fei Li, professor at Stanford University and head of AI Research at Google Cloud, hits the nail on the head: Solutions are designed by us humans. How we design them is critical to their success.

While the development of Big Data and neural networks offers us incredible opportunities to create solutions, the first question we have to ask ourselves is, which problem we intend to solve. Our approach to solution development is to focus on the user and his or her needs in accordance with our design philosophy. We want to understand what users are concerned with, what challenges they have, and how their world is changing. We supplement this design approach with the advantages of data and its exploration. It is not just a matter of bringing a design thinker together with a data scientist and thus putting together a multidisciplinary team that in itself offers significant added value. The aim is to promote multidisciplinary individuals who have their strengths in certain subject areas, one in design thinking, and the other in data science. Only by so doing can disciplines benefit from each other and jointly create solutions that optimize customer benefit. This gets clearer when we separate the views on a problem into the qualitative user view and quantitative data view. Design focuses on qualitative insights generated by observing the user, whereby data science entirely focuses on quantitative insights generated by exploring data. Joining both views leads to a holistic perspective. Our approach from problem identification to the completion of the optimal solution is based on the process defined by the Hasso-Plattner Institute of Design at Stanford [7] and is divided into five iterative phases.

Phase 1: Building empathy

User research is an integral part of a design-driven approach. The aim is to understand the user's environment, behavior, needs, and challenges. For this purpose, interviews are conducted and users are observed in their interaction with their environment and in their daily work. A broad understanding of the users' needs and challenges is built up, and hypotheses for satisfying these needs are defined and iteratively validated. This approach is fully user-centric and qualitative. To gain further understanding of user behavior, data from the user's environment can be collected and analyzed. With this quantitative data, the qualitative data from the research can be supplemented and analyzed over a large number of users.

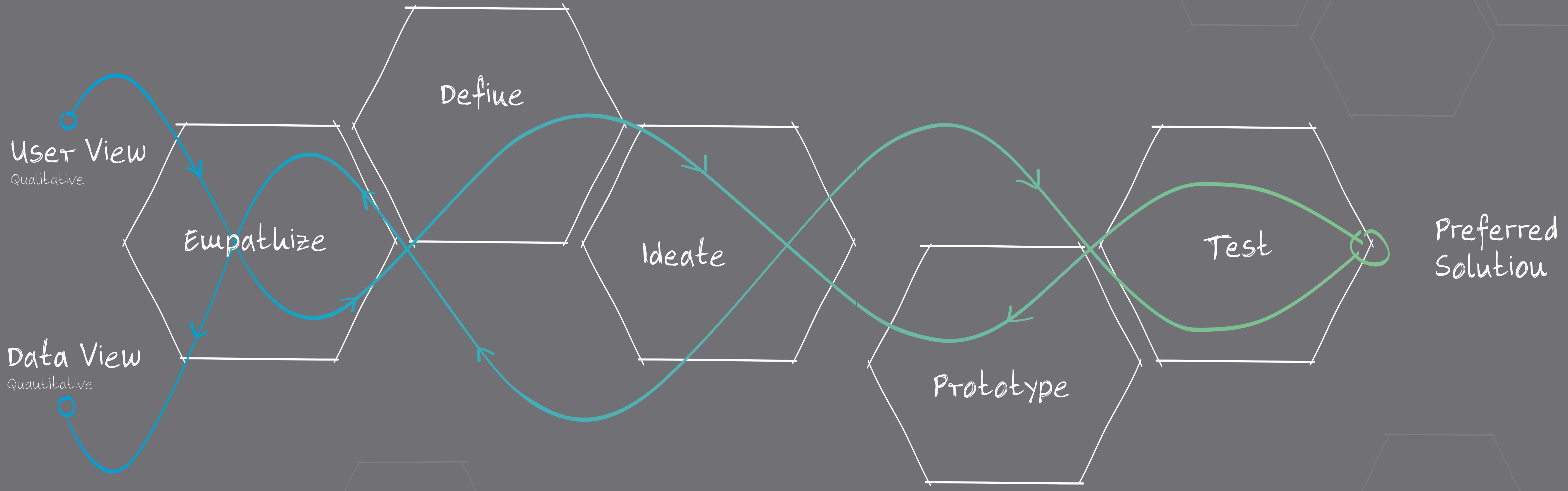
A simple example is that of a customer in a hip, metropolitan café. From observations of the design thinker, we find that the customer prefers black coffee in the morning "to wake up", but often buys a refreshing smoothie in the late afternoon. The data analyst also notes that black coffee is preferred in the morning and smoothies in the afternoon. He can't derive the reason like a design thinker, but instead determines the café's exact turnover and identifies the exact peak times. This is information the design thinker does not have. The findings from both views form a holistic picture, which leads to significant added value in the following phases.

Phase 2: Defining needs

This phase deals with the derivation of findings from the observations of the first phase. The aim is to formulate specific problems and objectives of the users from the rough problem definition rather than having to research specific issues. This is important because good solutions can only be found when a validated problem of our target group is solved.

In this phase, there is an unusually high demand for data analysis, which recognizes causalities and patterns and thus makes suitable deductions. At the same time, new questions can arise during the exploration of the collected data, which needs to be qualitatively validated using the design approach. Thus, the step of developing empathy and defining needs is carried out in constant iteration until clear problems and requirements can be defined. Turning back to our metropolitan café, data scientists could find a pattern regarding the peaks of the users entering the café, which correlates with the time table of the nearby subway. This pattern has to be qualitatively analyzed by the design team, if this is the case and why there is a remarkably high number of customers using a specific subway route.

User-Centered AI



Empathize: We want to thoroughly understand our users and identify their specific problems.

Define: By exploring patterns in data and validating hypotheses with user tests, we aim to define the core problems of our users.

Ideate: The extensive use of creative methods and the broadening of the solution space by introducing ML helps us to find innovative solution.

Prototype: Using the Lean Startup approach, we build prototypes, validate their value add and improve the solution iteratively.

Test: It must be ensured that the solution fully meets the user's need and model testing is important to validate the correct functioning of the AI.

Phase 3: Generating ideas

The task now is to develop suitable, innovative solutions for the defined problem. A wide range of creative methods can help. Solutions are found in close cooperation with users and relevant stakeholders. The involvement of data scientists at this point offers the opportunity to further open up the previously limited solution-space by introducing stakeholders to possibly unknown new and complex technologies such as AI. By means of impulse lectures and workshops, they can learn to correctly locate and integrate the latest technologies into their problem context.

The aim is not to convey an understanding of individual algorithms, but to create a basic understanding of the topic: What makes AI technology so interesting? What does machine learning mean? What possibilities does Supervised Learning offer? We, therefore, developed workshop formats to enable the user to understand AI and apply it to formulated problems. Furthermore, creative methods can help data scientists to creatively think about the use of data and algorithms to generate novel solutions.

Phase 4: Developing prototypes

If innovative solutions could be generated, it is important to implement the prioritized ideas prototypically during the development phase. Here we use the Lean Startup approach [8]. This approach aims to rapidly implement and test small product increments and hypotheses in close collaboration with users. The knowledge generated from these tests is fed back into further development of the product, which therefore evolves iteratively. By so doing, the user is actively involved in the development process.

The result of this procedure may not necessarily be a complete AI solution since data generation and preparation can take a long time. It is rather the identification of the optimal process and information through simple prototypes that make the solution tangible for users and relevant stakeholders. These prototypes should not be based on artificial data but should contain first findings as ascertained by the data scientists. The more detailed these prototypes are designed in the iterative process, the clearer the user requirements can be conveyed to the data scientists. Data scientists can then identify the optimal technical solution (AI, rule-based approach, etc.) within the framework of a technology benchmark and further develop it. The combination of design, Lean Startup and Data Science processes leads to rapid, target-oriented and sustainable solution development.

Phase 5: Testing solutions

Both in design (User Tests) and Data Science (Model Test), testing is of great importance. On the one hand, it must be ensured from the user's point of view that the solution fully meets the user's needs. On the other hand, model testing is essential to validate the correct functioning of the AI. The results of the tests lead to an iterative design of the product or even to the generation of new solutions. Iteration in the procedure is crucial, increasing the level of detail in the requirements specification, which in turn leads to optimally designed solutions.

What Were Our Experiences?

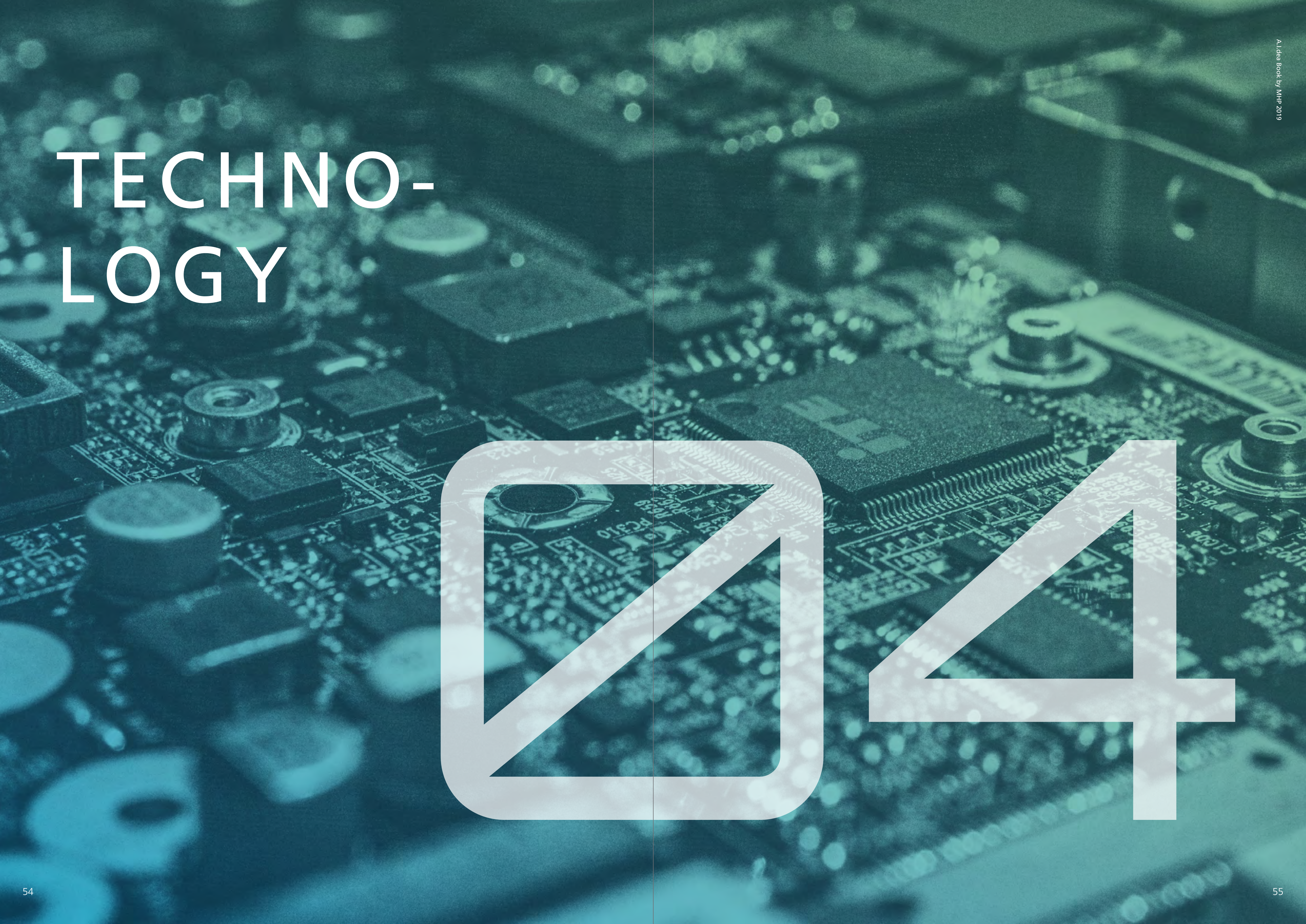
Within the realms of our projects, we were able to identify the following advantages of this approach:

- Resolution of real (specific) user problems.
- Development of the right solutions in a very short time.
- Focus on relevant information and functions.
- Generation of feasible solutions by early involvement of all relevant stakeholders ("fail before you start").
- Increased user acceptance of the developed solution.
- Realistic expectations and a strong commitment to the project by constant user-involvement.
- Unnecessary effort in data generation -cleaning and processing is avoided.

AI is a powerful toolset, which will lead to solutions we have never thought are possible. However, to find these solutions, we must understand the problems and the users affected thoroughly. Plus, the development should follow a lean startup approach by building minimal viable products, testing them with the users and improving them iteratively. If users are at the center of our approach, we will succeed in delivering exceptional products with high user acceptance and ultimately add value.

Despite its name, there is nothing 'artificial' about this technology – it is made by humans, intended to behave like humans, and affects humans. So if we want it to play a decisive role in tomorrow's world, it must be guided by human concerns.

Fei Fei Li | Professor at Stanford University and former Head of AI Research at Google Cloud



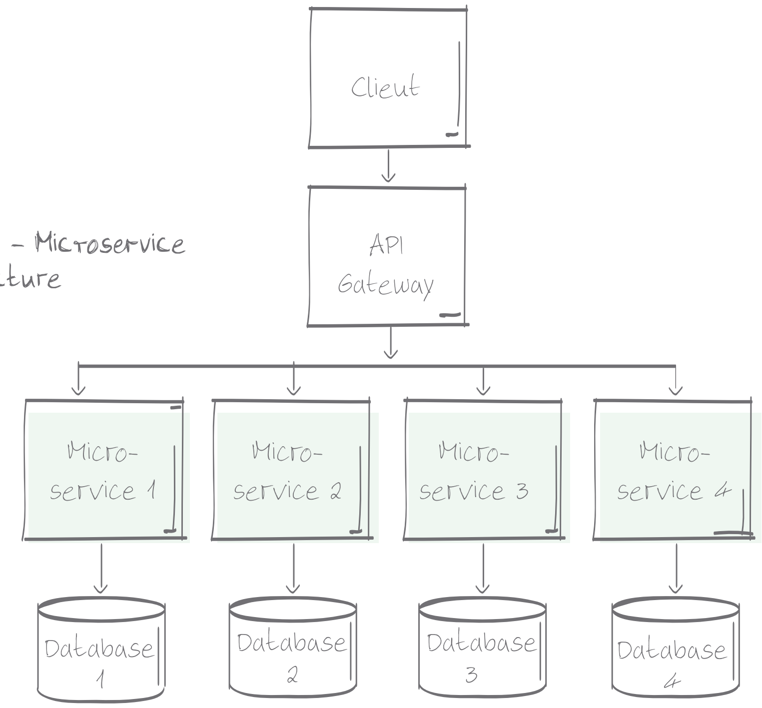
TECHNO-
LOGY

04

Containerized Machine Learning Architectures

By Felix Maximilian Roth & Fabian Wittke

Figure 1 – Microservice Architecture



Every day, an increasing number of Data Scientists work hard to develop and optimize models to solve machine learning (ML) problems that stretch the bounds of our imagination. Most of the time, this process starts on the data scientist's computer using a Jupyter Notebook as a means of rapid implementation, to get "a feel" for the data in hand. A fitting ML model can be found quite quickly, neglecting the need for a concept for its deployment. The consequence is a lack of a mechanism for accessing the model in any other environment.

Unfortunately, without a sound deployment concept and associated mechanisms, the best ML model has no real business value – It essentially remains a solution to a data science problem. Ideally, not only the model but also the required preprocessing steps should be made available for retraining purposes. Unfortunately, many AI/ML projects fail at precisely this point because it is here where knowledge of two very different disciplines is required: On the one hand, there is Data Science – rather new, explorative, with few standards and best practices. On the other hand, there is Software Engineering – well-established, with plentiful standards and accepted principles. Both of these disciplines are required for the successful implementation of a data-driven application. Therefore, a robust and scalable state-of-the-art solution is necessary for deploying ML models, using well-established architectural principles from software engineering, while taking potentially rapid changes implicit to the world of data science into account. In this article, we present MHP's approach to addressing this issue: Containerized Machine Learning Architectures.

Microservices Architecture

One of the most prominent architectural concepts in the time of Cloud Computing and the Internet of Things is microservices. In this concept, an application is defined as collection of small and independent entities – so-called services. Each service encapsulates only one sole responsibility, which is accessible through a well-defined interface. Additionally, a service potentially has its own database and can be independently implemented, extended, as well as deployed (see Figure 1 for an exemplary visualization of a microservice architecture). In combination with containerization of components, a microservices architecture provides many advantages for the implementation and provisioning of machine learning applications in an online environment, such as separation of individual process steps, dedicated deployment, scalability of individual services, and simple exchange of components.

Stream Processing

In such a scenario, services can be restructured, revised, and made available to the system as a new iteration through their design and deployment. But this does not apply to the ML model until now. The model itself is a static component in this construct, based on static historical data, which is fine for certain use cases with stable relationships. However, in today's software systems, very few components and processes are based solely on historical knowledge. Much more, the nature of systems is changing towards continuously improving systems. A large part of this change comes out of the Big Data world and is the introduction of the architectural idea of data streaming. Big Data established the value of insights derived from processing data which means in real-time (or streaming) applications that new data is continuously produced through capturing information, such as, user interactions, transactions, positions, and social media content, or also sensor data in an IoT (Internet of Things)-scenario, and processed on-the-fly. Stream processing in the field of Big Data does not only allow for detecting emerging insights, but it also enables continuous improvement of the used ML model.

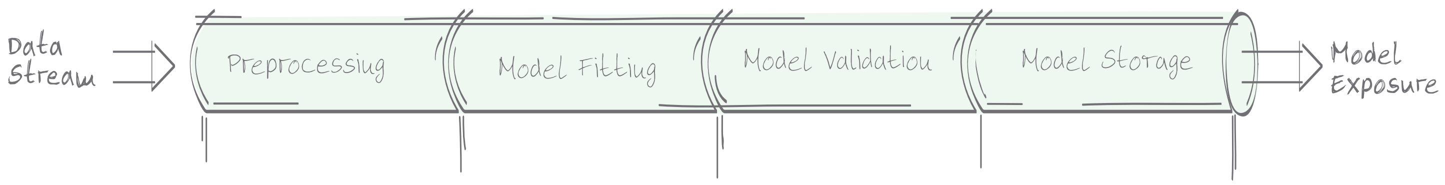
Continuous Adaptive Learning

This continuous improvement of the model can be achieved by establishing a retraining process based on the described permanent data input stream. Furthermore, streaming data makes continuous adaptive learning possible, which means a renewal of the used model through retraining, taking the newest data and insights into consideration to react more quickly to changing conditions – like the manner humans learn by observation of the environment. In order to enable such retraining, mechanisms are needed, which determine when a retraining of the model is appropriate on the input stream through metrics and quality checks.

Machine Learning Pipeline

The combination of the described architectural patterns with a continuous data input stream – providing a continuous process for model improvement – leads to the following ML pipeline (see Figure 2). With the initiation of retraining, the following standardized processes can be used to prepare the data and the training itself. Data arrives continuously at the first step of the pipeline where it is preprocessed, e.g., inconsistencies are handled, data is standardized, and features are extracted. In a second step, the model is trained or updated before several validations are performed to measure its quality. Key metrics include accuracy, precision, or recall, among others. If the model meets the conditions or outperforms the existing models, it is stored in a model storage, i.e., a database. A microservice can then retrieve the model and expose it via the service's well-defined interface so that other components can access the model. In this architecture, each step can be executed by a different microservice in a separate container, decreasing dependencies. This has several benefits; for example, technology-independent processing steps and modifications to the business logic of a functionality can be independently integrated by simple service replacement. Also, the ML model-specific tasks are decoupled from model provisioning.

Figure 2 - ML Pipeline



Parallelization and Scaling

Usually, in computer systems, the number of parallel users accessing a system correlates to the response time. Thus, more parallel users lead to higher response times. By using microservices, we can easily counteract this problem by the addition of further instances of the same service. A load balancer then distributes incoming requests among existing service instances. Thus, the response time decreases through parallelization of request processing. Moreover, service instances can be dynamically added and removed depending on the current load for saving resources and costs. With a microservice-based architecture, scaling is simple and, thus, response time is not necessarily an issue anymore. Furthermore, parallelization and scaling does not only help to reduce the response time for accessing the model but can be applied to every building block of the above-mentioned ML-pipeline as well to increase the performance.

Conclusion

This article shows that new insights gained from the work of data scientists only add business value in conjunction with established software engineering approaches. We have provided insights into an approach that we are currently implementing for several customers who recognize the necessity to adhere to stringent software engineering principles while letting data scientists continue to use their preferred development tools inside containers. Containerization, as well as stream processing, directly contribute to a scalable and maintainable pipeline for continuous improvement of ML models. We propose that in addition to these traits, frontrunners in the adoption of AI-technology are those who are currently implementing machine learning models as part of an end-to-end software system based on containerization and stream processing – those companies who see the value in bringing data science and software engineering together.

IoT Platform Interoperability

By Felix Maximilian Roth

An Enabler for AI

In the field of Artificial Intelligence (AI), machines try to mimic the cognitive capabilities of humans. AI usually makes use of machine learning mechanisms and, thus, requires a lot of data for training purposes. For such AI use cases, data is acquired through different data sources, e.g., existing records in Enterprise Databases or web crawlers. Nowadays, vast amounts of data are gathered through pervasive computing and the Internet of Things (IoT) – the ability of devices to interact with each other and support users in their daily tasks. It's not only possible to collect information on certain contexts with IoT, it even enables automatic alteration of these contexts for optimization purposes, e.g., by decreasing throughput of a production robot due to a predicted upcoming repair at the subsequent robot in the production chain, which requires some robust business logic. Consequently, the full potential of both IoT and AI can only be reached through their combination.

Currently, a lot of different IoT platforms exist, facilitating the development of such scenarios. Although they work well in research settings, several problems arise in real-world deployments where usually several IoT platforms coexist in the same smart space. These problems include, but are not limited to, heterogeneities in communication technologies, discovery mechanisms, interaction paradigms, data syntactics and semantics, application access, non-functional properties, as well as notification mechanisms. Indeed, these heterogeneities prevent interplatform interaction. Therefore, this article describes one solution for achieving interoperability between different IoT platforms to improve data availability in combined IoT-AI use cases and thus acting as an enabler to leveraging their full potential.

Enabling Interoperability Between Different IoT Platforms

Several interoperability approaches have been proposed in the past – each with its own strengths and weaknesses. The following section presents the XWARE approach [2][1] tackling the heterogeneities mentioned above and, further, working transparently, i.e., services are not aware of any mediating entity. Figure 1 shows the architecture of the framework. A communication component is responsible for handling interactions with actual services and applications with the help of plugins. A plugin defines the exact method of communication used by a specific platform and performs the transformation between the platform-specific and a uniform message format. Thus, a plugin is required for each platform to be supported. XWARE provides several skeleton classes and protocol implementations, accelerating the integration of new platforms.

The alignment component takes messages in the uniform format and processes them further. It employs the pipes and filters pattern [5], enabling great flexibility as well as extensibility through configuration-based context-aware filter assembly. Thus, messages pass through different filter sequences depending on the message's purpose. For example, a service registration message is passed to the discovery filter for converting service information and storing it in the service registry, whereas an application message traverses the service identifier filter, interaction filter, and application filter to translate the respective message layer one after another into the target format. For the transformation of message content, XWARE offers a simple automatic solution building upon extended WSDL files as well as a manual solution requiring code writing. The translated messages are then forwarded to the appropriate component – communication or notification.

The service management component keeps track of services joining and leaving the environment using an adjusted lease-based mechanism that allows for explicit (de-)registrations as well. The service management component stores each service in several service registries – one for each integrated platform – speeding up service transformations.

Last, the notification management component enables notification support among different platforms – even including platforms without native notification support to serve as producer. Furthermore, it translates between notification schemes, e.g., channels and subjects.

Abstractions facilitate such a framework by using uniform representations and patterns, e.g., service model, service discovery model, service access model, notification management model, or message model. Therefore, existing mechanisms need to be mapped to those abstractions. In this process, platform-specific information might get lost. Nevertheless, reducing the amount of required transformer specifications substantially, it seems to be a reasonable trade-off.

Evaluating the Solution

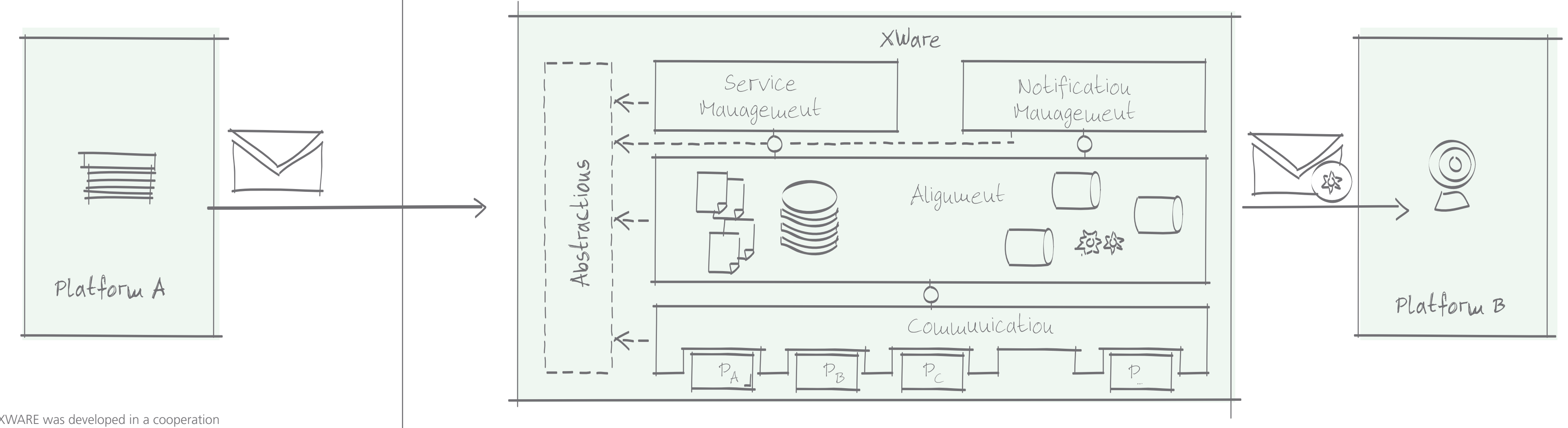
For evaluation, six different platforms have been integrated into the framework, and the prototype has been showcased in real-world scenario simulations of smart home automation. The overhead for integrating a new platform has been evaluated to be very appropriate, reaching from 24 to 1001 logical lines of code per plugin. As well, a cost evaluation has shown that the time for intercepting, translating, and forwarding messages is very short. Problems may arise if a synchronous RPC platform requests data from a publish-subscribe or tuple space platform due to the uncertainty when and if new data is produced and the requestor is blocked until a response is received. A cache component has been used in the evaluation to alleviate this problem. Overall, the overhead is more than acceptable considering the substantial benefits.

Conclusion

This article described the XWARE framework, which enables interoperability between different IoT platforms. Using various mechanisms to keep the framework customizable as well as extensible, it facilitates the integration of new platforms and services. The employed pipes and filters pattern even allows making changes to the transformation process, preparing for future changes.

Concludingly, XWARE (and IoT platform interoperability solutions in general) enhances service availability as well as accessibility and, therefore data availability. From an IoT perspective, AI provides the potential for robust automation mechanisms. Thus, at AI@MHP – which stands for End-to-End AI solutions – we believe that AI and IoT form a natural symbiosis.

Figure 1 – Architecture of the framework (cf. [1])



1) This article is based on [1],[2],[3],and [4]. XWARE was developed in a cooperation between the University of Mannheim and Grenoble University. The work was supported by the German Research Foundation (DFG) under grant BE 2498/9-1 and the Agence Nationale de la Recherche under grant ANR-16-CE92-0038-01

Panacea of AI

By Pooja Mukund, Almas Myrzatay

Advantages and Limitations of Deep Neural Networks

Across the board in Technology, Deep Neural Networks (DNN), are taking over in popularity. The name itself sets up a certain expectation as the word Neural comes from the structure being loosely based on the neuronal structure in the brain. DNNs are so regularly praised in the media, that their limitations are often overlooked. One study examining the main scholarly contributions to the topic of Neural Networks (NN) found that the advantages of NNs are more frequently published than disadvantages [1]. However, it is not only in literature where we see this trend. Every day, companies are using NNs to predict, classify, and cluster almost anything imaginable. While NNs have existed for years now, recently NNs have been making significant strides, which have contributed to their recent fame. For example, Google's Deep Mind AlphaZero Go was able to defeat the World Champion of Go showing Deep Learning is capable of learning strategy and intuition, much like a human mind [2]. IBM Watson also utilized Deep Learning to help diagnose patients with cancer and offer customized treatment plans by comparing patients' medical records to a large database containing inpatient and outpatient data [3]. With tech giants showing that the applications for Deep Learning are endless, it is clear to see the reasons for their rise in fame. NNs have infiltrated all fields of industry, but they have been commercialized to the point where it seems they are a panacea to all data-driven problems. While significant advancements in AI have often been linked to NNs, it is essential to recognize that there are both advantages and limitations in choosing and implementing an appropriate model.

To describe the advantages and disadvantages, we first need to understand how NNs work. NNs have three components: an input layer, a hidden layer, and an output layer (Figure 1). Deep NNs are NNs characterized by multiple hidden layers. These hidden layers play a key role in how NNs can outperform other Machine Learning models. Deep NNs are consistently praised for their lower prediction errors compared to other methods. The increased number of hidden layers leads to more learning, which then minimizes prediction errors. All NNs start with an input layer of data. In a DNN, data traverses from the input layer through multiple hidden layers to the output layer. The hidden layers are responsible for the pattern recognition capability in NNs. Each layer detects some hidden pattern from the previous layer. For example, in image recognition, one layer could detect edges, the second layer could detect shapes, and so on [4, 5]. The hidden layer's ability to detect patterns is what makes it so useful in applications with unstructured, unlabelled Big Data [6]. Additionally, NNs can be stacked together in various ways to create different forms of NN architectures. In practice, many types of NNs are coupled together in order to produce optimal results.

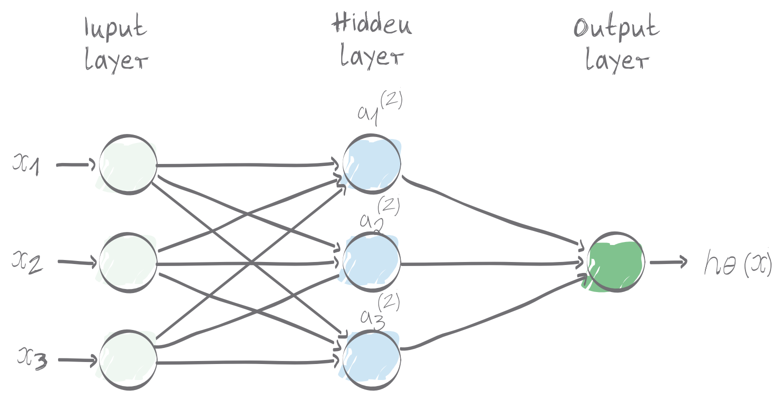


Figure 1 – ???

In the last decade, there has been a drastic increase in the volumes of data being produced in all warps of life. In the age of informatization, many companies have been pushing towards automation and digitalization. Low processing costs and storage have paved the way for a wide array of NN applications in speech, text, and image recognition. Previously, it was not possible to discern any pattern from this type of data, but Deep Learning changed that. Some researchers believe DNNs are intrinsically better than Single Layer NNs [7].

In contrast to traditional machine learning models, NN models can handle multi-dimensional features with high accuracy rate. Unlike training other previously established Machine Learning models where feature extraction must be performed to tell the model which variables are important, hidden layers in NNs are deduce the important features of the input without manual manipulation of the model. NNs offer a more robust approach to feature extraction due to the multi-layered architectural design [12]. Stacked NN layers paired with a large set of data yield high accuracy rates that often surpass any other supervised and unsupervised learning models. Each of the connections in a NN have a weight applied, which is adjusted in the training stage to decrease the model's total loss, a measure of the model's incorrect predictions. Then an activation function, either a linear function or non-linear function, is applied to a node to "activate that node" and determine the output of the NN. This allows for more flexibility when modeling the outcome.

Robust models make NNs suitable for a wide array of applications. Trained NN models on large datasets provide high accuracy rates to problems in many fields, including healthcare, as mentioned earlier, automotive, and consumer relationships. The application of data mining techniques for medical image classification reaches a classification accuracy of over 70% [14]. Research conducted by Itchhaporia et al. on cardiac images showed the average accuracy rate for NNs was 77%, while accuracy was 69% for junior residents, and 83% for staff radiologist [15]. Shayler et al. presented a NN use case in automotive engine management system as an alternative to 'lookup' tables. The research showed NNs capability to perform engine-fault diagnosis and performance optimization of electronic engine control systems [16]. In the consumer relation domain, NNs have been developed into conversational dialogue systems in the form of chatbots for customer support operations [17]. High accuracy rates of complex and wide-ranging applications of NN resulted in the popularity of NNs in business operations.

One of the main reasons NNs have seen a rise in popularity is due to their ability to "understand" Big Data. As companies start to collect more data, few technologies can handle these volumes and types of data as well as NNs. Studies show NNs perform better with a vast amount of "messy" data compared to a small amount of structured data [6]. Companies often want to use NNs due to their higher prediction rates, but what is often misunderstood is that these prediction rates can only be achieved with Big Data. IBM Watson has recently been criticized for giving doctors unsafe treatment recommendations for patients, but IBM claims that instead of feeding the model real patient data, doctors were providing hypothetical data, leading to inaccurate results emphasizing the need for large amounts of real data to be collected for the model to perform accurately. This is an issue commonly encountered when big NN architectures are not properly trained with sufficient amount of data, which leads to underfitting and subsequently leads to inaccurate results from the NN model.

For organizations that do not already have access to the infrastructure to collect Big Data, the process of initiating the collection of Big Data can be lengthy and expensive. Collecting Big Data requires the cost of an additional team of individuals who have the specialized skillset to implement Data Infrastructure as well as the cost of the hardware, software, and cloud infrastructure to be implemented. Furthermore, when data is collected, it is often not in a form ready for analysis, so there are additional costs associated with processing the raw data. Unlike many machine learning models, NN are capable of learning based on unlabeled data through the process known as Unsupervised Learning. However, NNs tend to perform better on labeled data, which can be expensive and time consuming to acquire. If not executed correctly, this could cause wasted staff time, incurring costs. While the benefits of NNs can be great, the costs also need to be considered, even though these costs may, in part, be indirect.

Another major disadvantage of the hidden layers in DNNs is that the learning period tends to be computationally expensive. By a method called Back-Propagation, weights between hidden layers are adjusted until output errors are minimized. By increasing the number of layers, the number of computations the algorithm must perform also increases, which can take time and computing power. DNNs can also be hard to train because earlier hidden layers take much longer to learn than later hidden layers [9]. In addition to the hidden layers being computationally expensive, NN can improve predictions by training and testing on larger amounts of data, but extracting, cleaning, and pre-processing the data can also be computationally expensive. The cost of NNs not only involves computational expenses but also hiring costs. Deep NNs are considered one of the most complex types of modeling techniques and, therefore, require people with a highly refined and specialized skillset. In addition to this, NNs require a lot of trial-and-error, so more time will need to be devoted to ensure the model is performing at a high accuracy rate. Moreover, trial-and-error process of optimizing NNs requires GPU powered infrastructure, which is costly, and requires specialized skills to setup and to operate. While feature

There are many ML applications, where the "why" matters more than the "what".

Why for example, was one particular product recommended to a customer rather than another? Why exactly was an image classified as containing an anomaly?

In many practical applications the black-box approach isn't good enough, and "explainability" must prevail over performance.

Johannes Keller | Head of Business Area Artificial Intelligence

extraction is not performed, the architecture of the Network needs to be tweaked until the optimal number of hidden layers has been found.

Despite significant breakthroughs in the many fields, NN's black box learning approach is not suitable for all use cases. Various factors such as data quality and quantity, ambiguity in model interpretation, and high financial overhead will influence model selection on a case by case scenario.

The amount of readily available and clean data is an important factor in the model selection process. Given small amounts of data, Support Vector Machine (SVMs) classifiers perform as well as NNs. For example, hybrids of SVMs and Hidden Markov Models (HMMs) that digitize handwritten digits perform better than Multilayer Perceptron (MLP) and Radial-Basis Functions (RBF) NNs [11, 13]. Besides, research by Jacobson et al. shows the performance of NN models is not significantly better than the performance by Decision Trees on "imprecise and uncertain" data [12]. Although NNs perform better with large data, classification models such as SVMs and Decision Trees perform equally well when data is scarce and imperfect.

As a non-parametric model, NN models lack interpretability and may require high operational overhead. This is because unlike other ML models like Regression or Decision Trees, NNs do not have interpretable features in the model that can explain how the output value was predicted. NN learning is black box learning approach, which cannot deal with uncertainty, and doesn't present direct relationships between input and output. The lack of relationship interpretability in NN models will require immense processing power to design the model, and experts to maintain it, which in turn increases operational costs. The lack of intuitive results provided by these models is arguably the main critique of Deep NNs. Some theories remain about how hidden layers produce certain outcomes, but much is still left to be explained [10]. In settings where interpretability cannot be sacrificed, other models are better suited. For example, many banks use non-neural network-based regression models for credit risk analysis in order to have indicators to explain to customers why they cannot get a particular loan. On the other hand, classification models such as SVM and Decision Trees can handle multi-class classification out of the box. Although, clean and processed data sets are prerequisites to good models, training NN models requires a specialized human skillset and expensive computational resources. On ther other hand, models such as Random Forest and Regression can be configured easily, they perform well on relatively small datasets, and have low training costs.

The ability to model data and predict outcomes has significantly advanced with the invention of DNNs. DNNs have risen in popularity due to their ability to effectively incorporate large amounts of messy and unstructured data with unprecedented accuracy. Despite these advancements, DNNs are not a universal solution. Like every machine learning model, however, DNNs come with their own set of limitations. These shortcomings include a lack of interpretability and high costs, both computationally and financially. Because of this, DNNs tend to be a better option when feature relationships are not a part of the question, when data is abundant and unstructured, and when the necessary resources and computing power to run the model are available. Many models and machine learning techniques have been developed in the areas of speech, text and image recognition. Although there is a myriad of models to select from, each technique comes with its own list of pros and cons. Depending on the use case, it is important to consider alternatives to NN and to keep in mind the economic feasibility associated with building a model.

From Text Data to Valuable Insights

By Yannick Fischer and Alexander Höweler

An Introduction to the MHP Adaptive NLP Pipeline

Looking at modern developments in data analysis, analyzing all variations of data has become possible and necessary, especially for unstructured data such as text data. Research has shown that only about 20% of all data is structured and available to enterprises, and the other 80% of the data is still unstructured, and most of it is text data. At MHP, we believe in the significant potential of unstructured texts and the meaningful insights, that can be derived from this source. However, leveraging the potential of raw text data and reaching a point where meaningful analysis is possible is not trivial and requires complex data processing.

For this purpose, we created the MHP adaptive Natural Language Processing (NLP) Pipeline. It offers a modular toolkit for preprocessing text data for subsequent, more in-depth analysis via NLP and is flexibly applicable for all purposes of text processing. That said, the pipeline is not a series of sequential processing steps but rather processing toolkit from which elements may be selected individually. The basis of all NLP analyses is the right **raw text**. This includes, but is not limited to, different sources such as news articles, social networking posts, comments, annual reports, etc. An example of such a text could be:

Raw text

"Yesterday, the weather was not good, but at least it wasn't raining".

First, the text is **cleaned** by lowering all letters, removing characters, and deleting unused spaces. Punctuations are kept since they are needed in the later Part of Speech tagging. This step is used to break the text down into a format, that is easily interpretable for machines, to simplify further processing steps.

Raw text → cleaned text

"yesterday the weather was not good, but at least it wasnt raining".

The subsequent step is the **tokenization** of the text. Here, the text or a set of texts are split into individual words to identify every single piece of text, to enable further processing steps, such as text mining or parsing with the extracted list of tokens.

Raw text → cleaned text → Tokenization

["yesterday", "the", "weather", "was", "not", "good", ",", "but", "at", "least", "it", "wasnt", "raining", "."]

Stop word filtering filters all stop-words, like "as, of, to, …", from the text in order to reduce the substance of the text or sentences to the meaningful core words. Note the nt from the negation of the verb wasnt: This step is done, as stop words have no substantial value to the NLP pipeline and for example, the required computing power or storage space can be reduced.

Raw text → cleaned text → Tokenization → Stop word filtering

["yesterday", "weather", "good", ", ", "nt", "raining"]

In **negation handling**, negations are identified for correctly interpreting the later sentiment, in other words the emotion, for keywords in the given text.

Raw text → cleaned text → Tokenization → negation handling

["yesterday", "the", "weather", "was", "not", "good", ",", "but", "at", "least", "it", "wasnt", "raining", "."]

Part **Of S**peech tagging allows us to identify the function of each word within a sentence. Verbs, nouns and adjectives are classified to further evaluate the role of each single word. The previously tokenized example shows how the algorithm detects the parts of the text:

Raw text → cleaned text → Tokenization → Part Of Speech tagging

"yesterday",	"the",	"weather",	"be",	"not",	"good",	",",
↓	↓	↓	↓	↓	↓	
NOUN	DET	NOUN	VERB	ADV	ADJ	

"but",	"at",	"least",	"it",	"be",	"not",	"rain",	"."]
↓	↓	↓	↓	↓	↓	↓	
CCON	ADP	ADJ	PRON	VERB	ADV	VERB	

Stemming and **Lemmatizing** identify the core of each specific word in order to associate the word with the correct meaning. After this step, the tokenized text looks as follows:

Raw text → cleaned text → Tokenization → Lemmatizing

["yesterday", "the", "weather", "be", "not", "good", ",", "but", "at", "least", "it", "be", "not", "rain", "."]

N-gram analysis is used to identify particular sequences of n items in the given text. This is applied in order to better classify the meaning of these sequences or isolate them from the text. The example shows the chosen word "weather" with an n-gram=1, meaning that to both sides of "weather", one word was added to the n-gram.

Raw text → N-gram analysis

["Yesterday", "weather", "was"]

Depending on the aim of the analysis, multiple texts can be **classified** to reveal emerged focus points. For classification, various approaches can be used: For example, dictionary based approaches, Deep Learning approaches, rule based methods, or combinations of all of these methods.

As mentioned above, the sentiment is the emotion that can be extracted from a text. This **sentiment** can be quantified as a **score**. This score can be calculated by using established scientific formulae or by a new metric individually developed by the analyst for the given topic. The higher or lower the score, the better or worse is the sentiment regarding the chosen topic. This helps to analyze whether the text has a positive or negative sentiment or something in between.

Raw text → cleaned text → Tokenization → Lemmatizing → Sentiment score

["yesterday", "the", "weather", "be", "not", "good", ",", "but", "at", "least", "it", "be", "not", "rain", "."]

→ Sentiment Score: -0,3 | slightly negative sentiment

As previously stated, MHP's adaptive NLP Pipeline can be used for a variety of use cases, that aim to analyze unstructured data. On the following pages, we present two NLP projects in which our Pipeline was used and adapted successfully.

PREDICTING STOCK PRICES WITH NLP

Analyzing Financial Markets using Sentiment Analysis

Background

To show the adaptivity of the MHP NLP pipeline, we applied it to the question "How easily quantifiable is the reaction of the stock market if political and monetary decision-makers raise their voice?" A central role in our approach is the real-time accessibility of sentiment in news articles. By having access to the sentiment carried in public media through an automated approach, better investment decisions can be made, as less time is spent to gather information, and more time can be used to evaluate it.

Approach

Figure 27 shows which components of the MHP adaptive NLP pipeline we use in our approach. Over 10,000 news articles from over 15 years are used for times where relevant announcements were made by political and monetary decision-makers, to estimate and predict the expected price of stocks associated with the U.S. real estate market. The news articles are processed by using parts of our adaptive NLP pipeline: Raw news articles are cleaned, tokenized and filtered for stop words; sentences relevant for specific topics are identified through POS tagging. Subsequently negations are handled, and sentences are classified individually. Using different concepts, ranging from machine learning to dictionary-based approaches, we estimate a sentiment score for every individual news article, allowing for a simple assessment of the influence of this article on stock prices.

Results

The results imply a significant influence of the public perception of statements made by decision-makers on financial markets. For every positively interpreted announcement by a decision-maker, the stock market reacts on average with an upward trend – and vice versa for a negatively perceived statement.

However, we did not only show the importance of sentiment carried in news articles for stock prices, but also the relevance of an automated approach making global sentiment accessible in real-time. Especially in a sector as volatile as the stock market, this enables organizations to improve their future investment strategies.

Our easy-to-use NLP approach allows for cross-sector use cases, for example, the real-time monitoring of global brand perception or the real-time identification of topics and the associated sentiment to react more appropriately to customers' needs.

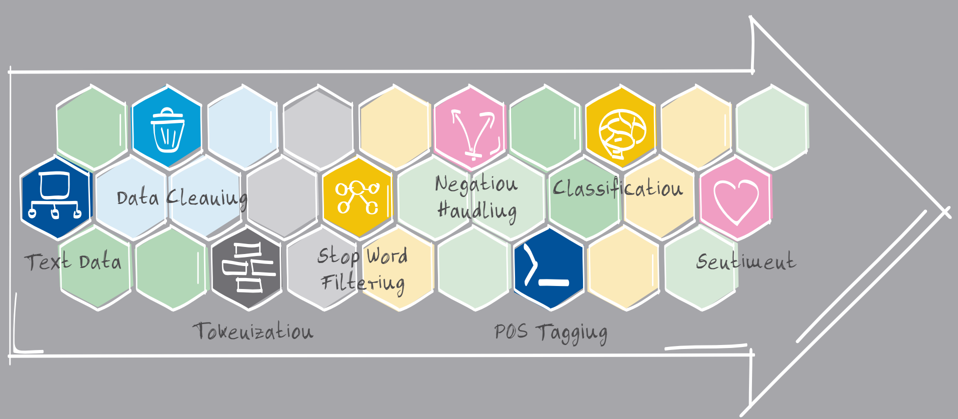


Figure 1 – MHP adaptive NLP Pipeline used for NLP-Based Stock Price Prediction

Figure 2 – Components of the adaptive NLP Pipeline used to gain customer insights

LISTEN TO YOUR CUSTOMER'S VOICE

Uncovering Customer Insights with NLP

Background

To highlight the broad applicability of MHP's adaptive NLP pipeline, the following section will demonstrate how to uncover customer insights with the use of NLP.

The honesty of customers giving feedback about using a product or a service can be one of the most valuable unstructured data sources. Hence, a profound analysis of such customer interaction data enables companies to improve processes based on customer needs and preferences. It also helps to find pain-points early and prevent them from gaining in significance. For this purpose, an NLP tool based on the above-mentioned MHP adaptive NLP pipeline was developed. We analyze the reaction of customers based on e-mails, regarding their customer journey, feelings about their product and service experience, their criticisms about pain-points or possible technical issues. The general aim is to classify e-mails regarding whether it contains customer points of complaint. We also discover what opinion a customer has and proactively reveal focal points well before they reach a critical mass.

Approach

Figure 2 shows which components of the MHP adaptive NLP pipeline were used for our approach. Predefined collections of words relevant to specific topics of interest – chosen by subject matter experts – constitute the base for our analysis and help to identify words that indicate a customer's dissatisfaction. In combination with POS tagging, we can analyze these words based on their position and the clause within a sentence. With the N-gram Analysis, certain parts of the e-mail are isolated to analyze their meaning more precisely. Other components of the pipeline are used to reduce the text's complexity.

Results

We are able to automatically label e-mails as customer complaints correctly with an accuracy of 95% - this in comparison to reading e-mails "manually". In combination with a suitable method of visualization, the manual workload of reading e-mails can be reduced significantly, providing hidden insights, originating directly from customer opinions.

RESPON-SIBILITY

05

First Do No Harm

By Heather A. Smith

Why AI Ethics Matter

Broad-scale use of Artificial Intelligence (AI) is in its early stages—few companies are ready to harness its power in a way that truly replicates human reasoning. As companies endeavor to use AI to deliver best-in-class services, it is important to consider the impact these technologies have on individuals and society. As we strive to deliver customer-centered experiences we need to understand how the data we collect and the knowledge we generate from AI can affect people and what unintended consequences stem from broad-scale AI adoption. How can companies deliver curated experiences and protect against potential threats? Are companies and consumers considering how AI can augment, rather than supplant human work?

In 2019 AI is becoming more mainstream. Companies like Google, Amazon, IBM, Microsoft, and Facebook are at the forefront of developing AI-powered products. These companies have vast resources to experiment. This gives them ample time to develop high performing products and solid go-to-market strategies. Smaller companies must be more tactical in their pursuit of AI because they have fewer resources and tighter margins. Tech titans are leading the way from a technological perspective, but how are they doing in terms of safeguarding consumers by following a code of ethics?

Google, IBM, and Microsoft have publicly available AI code of ethics statements [1, 2, 3]. Generally, they follow a similar precept: they outline their purpose in developing AI, list core principles, and describe what they will and will not pursue. Social benefits, privacy/data rights, accountability, and fairness are priorities cited by all three organizations.

AI as a Service

There is a key difference between greenfield AI projects like AlphaGo that is purely for innovation and using AI to be more agile and adaptive in business. AI as a service, refers to using artificial intelligence to deliver products and services more efficiently. Companies that effectively deploy AI-powered offerings will have a competitive edge in the years ahead—indeed, McKinsey estimates AI could generate an additional $13 trillion of value by 2030 and key global markets estimate its worth in the tens, if not hundreds of billions of dollars in 2020 [4]. Clearly, there is ample value to pursuing AI as a service.

To use AI to deliver customer-centered experiences, companies need a LOT of data. Collecting and storing that amount of data comes with risk. In today's world we must assume data breaches will occur. As we develop new technology-driven services, we cannot ignore our responsibility to protect the data we use to power our offerings. However, as we build, we need to carefully evaluate impacts and unintended consequences.

This article will examine the major ethical considerations in broad-scale use of AI as a service: privacy, dark use cases, and the human factor. To ensure consumer safety, companies using AI need to address ethical concerns. Ethical AI is more than protecting privacy—it is being able to foresee technology's influence. Most importantly, this article will offer some points of consideration for AI ethics.

Privacy

Most people value their privacy. According to the United Nations, it's a basic human right [5]. Customers trust companies to collect their data as a condition of using their services. Organizations using AI to enhance existing products and services (or deliver new ones) will collect vast amounts of consumer data—pictures of your face, voice recordings, shopping information, travel patterns, and communications data will all be gathered to create curated and seamless experiences.

However, that data can sometimes be used for purposes consumers did not initially consent to. This was a significant factor in the EU's move to protect people's data by passing the General Data Protection Regulation (GDPR). In the US, California recently enacted the California Consumer Protection Act (CCPA) to similarly safeguard consumer data. Clearly, privacy is a major consideration. Governments are acting to hold companies accountable as stewards of data, but is it enough?

According to Josefine Ehlers-Davidsen, Administrative Officer of Cyberawareness and Risk Assessment for Denmark's National Agency for IT & Learning, companies need to assume their data will be breached at some point. It is not a matter of if, it's a matter of when a breach will occur. Savvy companies will have mitigating strategies in place to minimize the damage to their customers. Davidsen also noted that it is vital to remember privacy is a highly individual concept that can be influenced by environmental factors and demographics. For example the implications for a lack of privacy will likely mean two different things for a person in the Nordics compared to a person in a country with a highly controlling state apparatus [6]. Who gets to decide what constitutes privacy? The solution is to ensure diversity and representation when developing privacy policies on AI, so that future applications of AI are built by the diverse many, and not the privileged few.

Dark Use Cases

While AI is promising, it also introduces levels of risk most organizations are not currently prepared to manage, no matter their size or budget. Too often, discussions of AI are limited to innovation and bringing products to market as quickly as possible. As practitioners, we need to discuss how to ensure our product is not only valuable, but how it addresses misuse or "dark use cases".

"Dark Use Cases" refer to unintended uses of technology products and processes. New technology changes the way we live our lives, and it also creates new opportunities for bad actors. As practitioners, how do we ensure the products and tools we build are not used for purposes we do not support? Like dark UI patterns, a dark use case creates a sub-optimal user experience. Where a dark UI pattern prompts the user of a product to act in a way that is not in their best interest as a condition of using the product, a dark use case can expose a user to unintended consequences of AI.

Politics & Social Media

A recent example of how ethics can make or break an organization comes from Facebook. Since the 2016 election, Facebook has come under intense scrutiny for failing to protect users from data theft and sentiment manipulation. Cambridge Analytica extracted vast amounts of Facebook user data without consent and used it to influence major voting campaigns and elections. Governments on both sides of the Atlantic Ocean want answers and are trying to understand what laws Cambridge Analytica broke or what loopholes they exploited when they created a Facebook app allowing them to harvest user information [7]. Facebook users were understandably upset to learn they were manipulated, and the company's public image suffered as a result.

Defense & AI

Google's AI ethical principles state they will not develop weapons or technologies that will harm people. However, they are currently working with the US Department of Defense—a move some employees are angry about. AI-powered weapons are a frightening proposition. Certainly, AI could lead to efficiencies in combat and possibly reduce risk to soldiers in the field, but what are the potential impacts on civilians? What if the technology is used to facilitate war crimes or human rights violations? Having a clear policy on weaponizing AI and adhering to it is vital to an ethical AI program.

Banking/Finance

Companies offering financial products are evaluating how AI can predict risk. Predicting which customers are most likely to default on loans or misuse credit can prevent loss of revenue. The idea of improving accuracy by employing AI to develop more sophisticated risk models is attractive to banks and creditors, but who decides what factors indicate a risky borrower? The financial services sector has already been shown to be pron to bias. The danger here is a future where existing prejudices are reinforced, and AI-powered risk models are used to unfairly discriminate. To ensure fairness, risk models need to be carefully examined for bias by a diverse selection of people.

Education

There are exciting partnerships between the UN and nonprofit organizations using AI to deliver education in Africa and South America. These projects aim to create joy around learning and deliver educational content in new and innovative ways to underserved communities. However, there are limitations to Western modes of data collection that have real consequences. Data's value can depend on cultural context. To make the best use of data to power AI in education and other industries, practitioners need to continually examine their biases and work to learn new ways of collecting and synthesizing data for AI development. One way to do this is to build inclusive teams and work closely with user groups and communities who make up target markets or who may be impacted by AI as a service.

Healthcare

What if your doctor could intervene before a major health problem irreversibly damaged your body? A predictive healthcare model powered by AI would provide multiple benefits. Medical and nursing schools could offer curriculum and specialized courses based on forecasted needs, pharmaceutical companies could develop drugs more efficiently, and humans could enjoy healthier lives. Unfortunately, AI could be used to penalize those who are deemed "high risk". Use of AI to evaluate patients could lead to a scenario where the most vulnerable populations are denied care. However, there is also a lot of promise. Using AI to augment human judgement could enable healthcare providers to make faster, more accurate diagnoses.

Mobility

AI as a service will have far-reaching impacts on mobility. Predictive maintenance, predictive supply chain, and connected cars/transit will change how we commute. Rather than reacting, consumers can enjoy a more trouble-free experience. Cars and public transit could use trip data to predict where we will go and automatically set navigation to our destination, avoiding traffic while recommending products and services. However, that technology could be used for surveillance. Technology practitioners and lawmakers will need to protect consumer privacy. Companies stand to gain significant efficiencies using AI to increase speed and accuracy and consumers will benefit from more trouble-free experiences. Communities can use AI to manage traffic and strategic development of transportation infrastructure.

The Human Factor

AI as a service has the power to disrupt entire industries. One of the most frequently cited aspects of this is job replacement. It is very likely that entire professions could be eliminated by AI-powered automation. No doubt, AI will change the nature of work. It is important for organizations to plan for this, so customers and communities are not adversely affected. AI should complement human experience, not eliminate it. Companies that address the human factor will enjoy greater credibility. AI can be used to enhance human capabilities and while it will change how we live and work, AI need not be seen as a threat. Rather, AI/human partnership should be viewed as a "new division of labor" [8].

Ethics is a vital consideration for any organization looking to offer AI as a service. This cannot be an afterthought—ethics must be part of product development, service delivery, and after-sales. Companies that fail to properly care for their customers' data will face fines, scrutiny, and loss of credibility.

AI has tremendous potential to transform businesses and individual lives. However, there are significant risks. Recently Elon Musk and Alibaba founder, Jack Ma argued about whether AI will do more harm than good. Just days later, several tech giants met in Switzerland to discuss the urgent need for a code of AI ethics. However, there seems to be a reluctance to propose actionable behaviors.

For companies to practice ethical AI, they must:

- Develop a position on why the organization develops AI-powered products and services that includes clear value statement on what the organization will and will not do
- Assemble a diverse committee of key stakeholders to mitigate bias and more accurately assess potential threats/dark use cases
- Have clear policies on AI as service that address consumer protection
- Follow the policies and periodically evaluate AI policy against current conditions

It is not enough to offer best-in-class experiences—companies must demonstrate commitment to ethical behavior for their consumers and society. AI as a service will change how people work, travel, live, and learn. Accountability is not negotiable.

Figure 1 – Sustainability Development Goals [2]

17 Sustainability Development Goals as a blueprint to achieve a better and more sustainable future for all.

Why Sustainability and Profitability Are Not Contradictory

By Benedikt Bothur & Felix Maximilian Roth

Sea levels are rising, the Amazon – the lung of our world – is literally burning, and longer, more intense droughts threaten crops, wildlife, and freshwater supplies. These omnipresent issues are currently being covered by news across the globe. To tackle climate change, which according to the UN Secretary General António Guterres is "the biggest threat to global economy" [1], much attention has been focused towards the replacement of old technologies with advanced and innovative solutions, mostly part of a variety of new "green" production initiatives. While being very enticing, the rapid implementation of new technologies across all industries is somewhat unrealistic. Particularly heavy industries have a substantial impact on the consumption of resources and subsequently are a mass generator of waste and emissions. Although efficiency improvements of industrial processes are pivotal to the world's sustainable future, businesses are forced to generate short-term profit for their very own survival rather than making investments into a sustainable future. To reduce a company's environmental impact and develop a profitable but sustainable business, why not apply an already existing set of appropriate technologies that will improve efficiency across existing processes and structures? If approached in the right manner, we firmly believe that Artificial Intelligence (AI) cannot only serve as a driver for the world economy but has the potential to make a positive social and environmental impact.

To clarify how we envisage ways in which this can be achieved, we first need to look at the Sustainability Development Goals (SDGs), proposed by the United Nations (UN). Then, several areas for improvement will be presented, before analyzing AI use cases in these areas for individual merits and potentials regarding sustainability and profitability, and their contribution to the SDGs.

Sustainability Development Goals

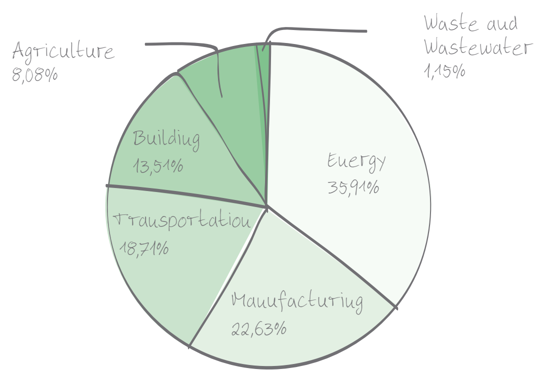
In 2015, the United Nations developed the 17 Sustainability Development Goals (SDGs) [2] as a "blueprint to achieve a better and more sustainable future for all" [3]. These goals aim at sustainable development, which promotes prosperity while protecting the environment (see Figure 1). Only by adhering to these goals can we tackle climate change and preserve an intact natural environment that is inextricably intertwined with the economic basis of every company. The SDGs demand action, innovation, and the urge to drive change. AI will not only support this change, but it will be one of the fundamental factors.

Having understood the commonly recognized and accepted goals for establishing a sustainable world, the next section briefly introduces different sectors where we see great potential for improvement.

Sectors with Sustainability Potential

We see major fields of improvement emerging through AI in sectors producing a lot of waste or greenhouse gases (GHGs) and sectors with excessive use of valuable resources. This includes, but is not limited to, the following sectors, ordered by GHG emissions in Germany (see Figure 2): energy, manufacturing, transportation, building, agriculture, waste & wastewater.

Figure 2 – German GHG Emissions Share by Sector in 2018 [4]



The following quoted figures apply to Germany in 2018 [4]. The interested reader is referred to [5], where similar figures are presented for the United States in 2017 from the United States Environmental Protection Agency. A comparison supports the assumption that these figures are representative of other leading industrialized countries.

The most significant contributor to greenhouse gas emissions is the energy sector, which was responsible for more than a third of GHG emissions in Germany in the year 2018. Fortunately, a declining trend is already visible, with a 2% reduction since the year 2017. The sector includes companies producing and selling energy. Henceforth, in our quest to find AI use cases, we will neglect fossil and nuclear energy sources since the aim should be their termination. Therefore, we will only consider the renewable energy sector. Especially in energy distribution and management, AI seems to be a promising candidate when it comes to improving supply and demand forecast as well as the management of smart grids for reliable and clean energy provisioning.

The manufacturing sector, which accounted for 22% of GHG emissions both in the US and in Germany in 2017, includes all kinds of production industries and thus has several overlaps with other mentioned sectors. Manufacturing processes are defined as the set of processes to transform raw material into the desired size and shape with required properties and characteristics. Companies operating in this sector can use data and AI in conjunction with each other to reduce emissions in this industry sector through optimization of production processes as well as significantly decreasing wastage (see, for example, our MHP A.I.deaBook articles on production efficiency and quality control). The transportation sector, responsible for 29% of GHG emissions in the US in 2017 but 18% in Germany, comprises corporations that provide infrastructure and services to move people and goods. Thus, there are four components required for these services: infrastructure (e.g., roads), vehicles (e.g., buses), operations (e.g., traffic management), and users (e.g., passengers). Especially in operations and infrastructure, we see massive potential for AI. Through, e.g., the application of genetic algorithms to routes or cargo, AI can contribute to a safer, more efficient, and more sustainable transportation industry [2].

Emissions in the commercial and residential or building sector mainly arise from the burning of fossil fuels for heat, electricity, and other in-house operations. Between the years 2017 and 2018, a reduction of 1% to 13.5% is expected for Germany. The type of power plant and the fuel used are decisive for the emission load attributable to a kilowatt-hour of electricity. Another critical factor is the degree of building insulation so that there is a direct correlation to the outdoor temperature – especially in winter. AI already plays a crucial role in smart buildings, by helping to turn raw data into actionable insights. For example, Forbes describes that with AI the operational efficiency and utilization of assets can be ensured, and the comfort level of occupants can be improved [6].

In the agricultural sector, which contributed almost one-tenth of GHG emissions in the US and Germany in 2017, farmers cultivate soil, produce crops, and raise livestock. On the one hand, approximately 11% of the world's population is currently suffering from hunger worldwide [7]. On the other hand, the Bill and Melinda Gates Foundation mention a population increase in Africa's – and thus the world's [8] – poorest countries [9], which are most susceptible to hunger and famine, meaning that in absolute terms, an increase in people suffering from starvation is to be expected. It requires massive increases in productivity and efficiency of agricultural activities to counteract the aforementioned trend. Through monitoring and managing yield, soil, and livestock, the agricultural sector can indeed benefit from AI to reach this goal. Furthermore, the food sector as part of agriculture is responsible for tons of food and packaging waste. Whereas transportation-related issues could be addressed as previously discussed (see transportation sector), problems caused by excessive packaging and the sheer demand of food can, in part be solved by AI-optimization of demand forecasting algorithms – This we have already implemented for a major food manufacturer, for example.

With an increasing world population [10], drinking water is getting scarce, while wastewater is increasing. In 2017, 29% of the world's population did not even have access to clean drinking water on-premises [11]. To improve water management and supply, AI can help with water demand prediction as well as infrastructure enhancement and maintenance,

AI Techniques

- **Machine Learning** (ML) includes supervised and unsupervised machine learning techniques, deep learning, as well as traditional statistical approaches.
- **Natural Language Processing** (NLP) tries to train computers comprehend natural language.
- **Computer Vision** (CV) is a field that attempts to teach computers understand digital images or videos.
- **Audio Processing** (AP) tries to interpret acoustic information correctly.

e.g., for leakage identification or even leakage prediction. Another vital issue is waste and its management. Inappropriate waste management can be the cause of epidemics and parasites, as well as the contamination of groundwater and soil [12]. Furthermore, waste disposal generates large amounts of GHGs, e.g., through the incineration of waste. For instance, AI can help to improve automated sorting processes for disposal and recycling decisions [13] – Visual Inspection and anomaly detection systems, as shown in our A.I.dea Book article "Automated Visual Inspection" can be easily adapted to perform these types of tasks.

In many of the cases mentioned above, a combination of AI with Internet of Things (IoT) technologies can be hugely beneficial and is, in parts, mandatory for data ingestion or automation – Also, see article "IoT Platform Interoperability – An Enabler for AI". This is because in many of these use cases, data has to be gathered via sensors, and changes to the environment need to be effectuated via actuators. So, for AI to be effective in reaching SDGs, many use cases rely on the close interaction of these two fields (IoT and AI).

Sustainability Versus Profitability?

This section presents a non-exhaustive overview (see Table 1) on use cases for the sectors mentioned above including indicatory AI-driven effects concerning sustainability and profitability. Furthermore, we indicate potential AI techniques we deem best suited for these use cases, as well as the SDGs addressed by these scenarios.

To comply with the goal of affordable and clean energy (SDG 7), the energy sector must minimize the use of fossil fuels. The technologies currently available, based on renewable energy sources (e.g. solar, wind, hydroelectric), show great potential in the transition to a greener energy sector. However, renewable energy sources are unreliable due to their intermittency. On cloudy or windless days, dips in renewable energy supply are currently smoothed out by backup energy sources such as coal-fired power plants or diesel generators. AI has the potential to pave the way to a low carbon electricity grid by rapidly processing, analyzing, interpreting, and especially acting upon the flood of data the energy sector generates. In such a system, energy from renewable sources can be stored locally and retrieved quickly during periods of low supply in renewable energy generation. Operating a smart grid does not only require a forecast for the electricity demand, production, or weather; it also needs to provide constant monitoring to ensure the electricity supply meets demand. Moreover, the smooth management of the varying contributions from renewables, nuclear and fossil fuel energy in the system is a key aspect. Such a smart grid complies with SDG 7 (Affordable & Clean Energy), SDG 11 (Sustainable Cities and Communities) and SDG 12 (Responsible Consumption & Production). Manufacturing and supply chain has been under the "optimization microscope" since the industrial revolution, and this is not destined to change any time soon. Sustainable manufacturing minimizes adverse environmental impacts while conserving energy and natural resources [14]. One use case with great potential is automated quality checks of objects as part of a real-time visual inspection system embedded in a manufacturing environment (see, for example, MHP's A.I.dea Book article "Automated Visual Inspection"). This enables the early detection of defects and leads to cost-savings in production due to higher efficiency while increasing product quality. For instance, we recently implemented a computer vision system for an optimized molding procedure resulting in reject reduction at a major German automotive supplier (SDG 9, Industry, Innovation, and Infrastructure). In general, applying a sustainable production solution (SDG 12, Responsible Consumption & Production) is a win-win because AI delivers improvements without high up-front costs and often, without changes to underlying processes.

In the transportation sector, several use cases have been identified that drive sustainability and profit at the same time. Predictive maintenance of infrastructure, e.g., roads, has the potential to increase infrastructure quality and safety while cutting costs. Usually, sensors are used to gather data on the infrastructure. Those sensors can be, e.g., cameras, audio sensors, or motion sensors. Using an model from machine learning, audio processing, or computer vision, the collected information is then processed and classified and can be used for planning and decision making. For example, Vialytics [15], a German startup, developed a computer vision-based system that scans road conditions and extracts maintenance measures. The outcome can be used to efficiently initiate remediation to extend the service life of the road. Predictive maintenance influences SDG 9 (Industry, Innovation, and Infrastructure) and SDG 11 (Sustainable Cities and Communities).

AI-supported traffic management can contribute to emission reduction and, from a macro-economic perspective, also reduce costs through saved time. Cameras and other sensors harvest data on traffic flow, emissions, and weather, to name but a few. This data can then be used with machine learning, computer vision, or reinforcement learning to optimize the traffic, based on predefined criteria. In addition to that, fleet, route, and cargo management may reduce emissions, resource usage (e.g., vehicles or employees), and, therefore, costs. Such management enhancements address SDG 9 (Industry, Innovation and Infrastructure) and SDG 11 (Sustainable Cities and Communities).

The building sector benefits significantly from advances in AI through an efficiency increase of used resources. Deepmind, a subsidiary of Google, uses machine learning to predict peaks of their data center's energy demands and optimize cooling as required. Consequently, their energy use has been reduced by 40 percent [16]. Integrated automation systems in buildings improve energy efficiency by analyzing data from sensors embedded in, e.g. doors, windows or air conditioning systems. In addition, optimized and personalized living and working environments are the result. Another interesting aspect is AI-enhanced site location. Joseph Sirosh, corporate vice president of Artificial Intelligence and research at Microsoft, recently stated: "Geographic informa-

POTENTIAL ANALYSIS

(ML: Machine Learning, AP: Audio Processing, CV: Computer Vision, NLP: Natural Language Processing)

12

	Use Cases	AI-driven Effects	Potential AI Techniques	Addressed SDGs
Energy	Smart Grid	↘ Costs ↘ Air Pollution ↗ Reliability	ML	SDG 7 SDG 11 SDG 12
	Predictive Maintenance of Infrastructure	↘ Costs ↘ Downtime ↗ Efficiency	ML AP CV	SDG 7 SDG 12
Manufac-turing	Quality Control	↘ Waste ↘ Costs ↗ Efficiency	ML AP CV	SDG 9 SDG 12
	Simulation and Manufacturing Planning	↘ Waste ↘ Costs ↘ Risk	ML	SDG 9 SDG 12
Trans-porta-tion	Predictive Maintenance of Infrastructure	↘ Costs ↗ Efficiency ↗ Safety	ML AP CV	SDG 9 SDG 11
	Traffic Management	↘ Emissions ↘ Costs (from a macro-economic perspective) ↗ Safety	ML CV	SDG 9 SDG 11
	Fleet, Route, Cargo Management	↘ Resource Usage ↘ Costs ↗ Utilization	ML CV	SDG 9 SDG 11
Buildings	(Energy) Demand Forecast	↘ Costs / Energy Usage ↘ Emissions ↗ Efficiency	ML CV	SDG 11 SDG 12
	Site Location	↘ Risks ↘ Costs	ML NLP CV	SDG 11
Agricul-ture	Data-driven Farming	↘ Resource Usage ↗ Yield ↗ Quality ↗ Profit	ML AP CV	SDG 2 SDG 12
	Enhanced Demand Forecast	↘ Waste ↘ Costs ↗ Product Availability	ML NLP	SDG 2 SDG 12
	Food (Pre-) Processing	↘ Waste ↘ Costs ↗ Efficiency	ML AP CV	SDG 2
Waste & Waste-water	Predictive Maintenance of Infrastructure	↘ Costs ↗ Efficiency ↗ Safety	ML AP CV	SDG 6 SDG 11
	Pump Station Management	↘ Energy Usage ↘ Costs	ML	SDG 6 SDG 11
	Waste Sorting	↘ Costs ↗ Recycling ↗ Safety	ML CV	SDG 6 SDG 11

tion systems [GIS], which can correlate and analyze location in time and space and integrate it with many other types of information—and then serve it up for higher-order AI to be applied on it—are particularly interesting" [17]. This can result in an optimized location for a wind farm or solar panels but can also be used for any type of building, e.g. wind-optimized building orientation. With the right combination of machine learning, computer vision, and natural language processing for the extraction of information, for example in building plans, AI can play a crucial factor in the building sector, thus enabling sustainable cities (SDG 11 Sustainable Cities and Communities) and responsible production (SDG 12 Responsible Consumption and Production).

In agriculture, the relationship between profitability and sustainability is most apparent: With the help of advanced techniques, companies not only increase their revenue by speeding up the production process through automation, maintenance time reduction, and hence the production downtime, but also deliver an excellent customer experience by predicting their likes, dislikes, and desires which leads to increased profitability through reduced wastage. Data-driven farming reduces resource usage (e.g. water, pesticides, etc.) while increasing both yield and food quality which in consequence increases profits. Usually in these smart farming scenarios, sensors are distributed around the farm with varying degrees of complexity. For example, sensors can be placed on a per-field basis to measure the soil or per plant/animal to obtain information at a much finer granularity. Complex models from machine learning, audio processing, or computer vision can, among others, then be utilized to optimize irrigation or feeding cycles. Besides, the data-driven farming use case is very generic and includes a multitude of smaller use cases, for example, yield management, pest management, or irrigation management.

Picture the yields from the data-driven farm described above: Sorting products or raw materials constitutes a time-consuming process. TOMRA, a leading sorting and collection solutions provider in this sector, uses a unique machine learning algorithm to analyze different aspects of fruits or vegetables for sorting and identifying possible anomalies [18]. Additionally, the existence of digital data flows enables the prediction of customer demand. The analysis and search for patterns and correlations are mostly based on historical sales data. Applying AI-techniques opens up the world of multidimensional pattern recognition where correlations between customer behavior, weather changes, seasonal and regional conditions, promotional offers, and price changes from competitors can be explored. AI-enhanced supply chain management may even lead to optimized stock replenishment and improved accuracy by reducing forecasting errors by 20-50% [19]. Furthermore, lost sales due to products being out of stock can be reduced by up to 65%. Overall, AI-enhanced supply chain management results in a reduction of wasted supplies, an increase of (affordable) product availability, and product and market context-based replenishment. Thus, this contributes directly and indirectly to SDG 2 (Zero Hunger) and SDG 12 (Responsible Consumption & Production).

In the waste and waste-water sector, predictive maintenance of infrastructure (like in transportation) has great potential to increase infrastructure quality, and thus service- and product quality and safety, and decrease maintenance costs. Water is an increasingly scarce and valuable resource, and according to Yaron Dycian, chief product and strategy officer for WINT Water Intelligence, water damage is the "silent killer" [20]. The company uses AI to detect and stop leaks in water distribution, thus reducing water consumption and preventing damage and commercial and government buildings. "The moment it [editor's note: WINT's AI-based solution] gets deployed", Dycian explains, "each device starts using machine learning to understand and analyze the local flow patterns, which will differ from location to location and season to season"[20]. After a learning phase, the normal patterns are established, and the implemented system can detect anomalies, pinpointing the exact location and nature of a leak. Various AI techniques are applied for this use case, which helps to achieve SDG 6 (Clean Water and Sanitation) and SDG 11 (Sustainable Cities and Communities). AI for waste management plays a crucial role in SDG 11 (Sustainable Cities and Communities). Oscar, a zero-waste AI solution by Intuitive [21], encourages users to recycle correctly by recognizing various items in the hands of users and suggesting the most appropriate means of waste disposal.

Conclusion

In summary, Artificial Intelligence provides a vast set of powerful methods for enabling or accelerating progress towards achieving the Sustainable Development Goals, by applying techniques that bring improvements mostly without adversely affecting established industrial processes that are pivotal to the world's sustainable future. Whether we are beneficial for our planet or harm it is ultimately not a question of whether or not we employ intelligent machines, but how we utilize and deploy them – hence, what we ask them to do for us. Asking machines to assist us in caring for our planet is an important step for climate change and environmental issues.

Stephen Badger, chairman of Mars Inc., summarized the point we are making when he explained his perspective as a major player for consumer food: "Addressing climate change or not addressing climate change will lead to an increase in raw material pricing" [22]. He furthermore outlined: "The magnitude of issues facing the planet is such that business has got to be front and center in that dialogue and in addressing those issues" [22]. Consequently, operating in a sustainable manner ultimately means a surplus in long-term profitability. It's time for companies to actively take action – not merely to accelerate the transformation toward becoming a more sustainable enterprise, but also as a means of becoming more profitable. At MHP@AI, we are not only considering classic KPIs, but are starting to include sustainability KPIs, such as efficient resource usage, energy savings, and waste savings.

Maybe it's time to include AI in your sustainability roadmap? We are more than happy to help.

FURTHER INFOR- MATION

06

REFERENCES

Chapter Quality 1.1

[1] V. Chandola, A.Banerjee, V. Kumar: "Anomaly Detection: A Survey". In ACM Computing Surveys, 2009.

[2] P. Bergmann, S. Löwe, M. Fauser, D. Sattlegger, C. Steger. Improving Unsupervised Defect Segmentation by Applying Structural Similarity to Autoencoders. In arXiv preprint arXiv:1807.02011 (2018)

[3] Ebrahimzadeh, Reza & Jampour, Mahdi. (2014). Efficient Handwritten Digit Recognition based on Histogram of Oriented Gradients and SVM. International Journal of Computer Applications. 104. 10-13. 10.5120/18229-9167

[4] Kubernetes Homepage, Concepts – Pods. Retrieved August 29, 2019, from https://kubernetes.io/docs/concepts/workloads/pods/pod/s", in Proceedings of the 34th International Conference on Machine Learning, 2017.

Chapter Quality 1.2

[1] Lamba, H. (2019). One Shot Learning with Siamese Networks using Keras. [Blog] Towards Data Science. Available at: https://towardsdatascience.com/one-shot-learning-with-siamese-networks-using-keras-17f34e75bb3d [Accessed 6 Sep. 2019].

Chapter Quality 1.3

[1] https://edition.cnn.com/2019/11/13/tech/facebook-fake-accounts/index.html

[2] Jenelle Bray: Automated Fake Account Detection at LinkedIn, 2018 https://engineering.linkedin.com/blog/2018/09/automated-fake-account-detection-at-linkedin

Chapter Efficiency 2.1

[1] Waters, R. and L. Urquhart. 2019. "World Preview 2019: Outlook to 2024". EvaluatePharma®: 12th edition. June 2019.

[2] Wang, C. and X. B. Liu. 2013. "Integrated production planning and control: A multi-objective optimization model". Pp. 815–830 in Journal of Industrial Engineering and Management. Vol. 6, no. 4. Barcelona: OmniaScience.

[3] P. W. Koh and P. Liang, "Understanding Black-box Predictions via Influence Functions", in Proceedings of the 34th International Conference on Machine Learning, 2017.

Chapter Methodology 3.1

[1] TechRepublic (2017): 85% of big data projects fail https://www.techrepublic.com/article/85-of-big-data-projects-fail-but-your-developers-can-help-yours-succeed/

[2] CDQ AG, HEC Lausanne (2019): Managing data as an asset https://www.cc-cdq.ch/news/whitepaper-managing-data-as-an-asset

Chapter Methodology 3.2

[1] MIT Sloan Management, BCG Henderson (2019): How to win with Artificial Intelligence https://www.bcg.com/de-de/publications/2019/how-to-win-with-artificial-intelligence-ai.aspx

[2] IT Production Online (2019): Closed-Loop Manufacturing – MES Funktionen als Bausteine einer IIoT-Plattform https://www.it-production.com/industrie-4-0-iot/closed-loop-manufacturing/2/

[3] Porsche Consulting, MHP (2019): Automation of Production Planning Enhanced by AI https://www.mhp.com/fileadmin/www.mhp.com/assets/images/ecosystems/MHPWhitepaper_Manufacturing.pdf

[4] CA Technologies (2014): Continuous Delivery – The ING Story: Improving time to market with DevOps and Continuous Deliver https://www.slideshare.net/CAinc/continuous-delivery-the-ing-story-improving-time-to-market-with-devops-and-continuous-delivery

[5] Porsche Consulting (2019): Leading the Way to an AI-driven Organization https://newsroom.porsche.com/dam/jcr:4139a15a-29d5-4207-af0e-6751e8c31eae/Leading%20the%20Way%20to%20an%20AI-Driven%20Organization_2019_C_Porsche%20Consulting.pdf

[5] TechRepublic (2017): 85% of big data projects fail, but your developers can help yours succeed https://www.techrepublic.com/article/85-of-big-data-projects-fail-but-your-developers-can-help-yours-succeed/

[7] Prediction Machines (2018): The Simple Economics of Artificial Intelligence.

[8] IT Revolution (2016): The DevOps Handbook: How to Create World-Class Agility, Reliability, and Security in Technology Organizations

[9] BARC (2018): BI Survey 2018 http://barc-research.com/bi-survey-18/the-bi-survey-18-the-results/

[10] MIT Sloan Management Review (2018): Don't Let Artificial Intelligence Supercharge Bad Processes https://sloanreview.mit.edu/article/dont-let-artificial-intelligence-supercharge-bad-processes/

Chapter Methodology 3.3

[1] E. Strickland, "How IBM Watson Overpromised and Underdelivered on AI Health Care", 2019. [Online]. Available: https://spectrum.ieee.org/biomedical/diagnostics/how-ibm-watson-overpromised-and-underdelivered-on-ai-health-care.

[2] IBM, "IBM Watson TV Commercial, 'Watson at Work: Healthcare'", 2019. [Online]. Available: https://www.ispot.tv/ad/wk_G/ibm-watson-watson-at-work-healthcare.

[3] M. Korolov, "6 reasons why AI projects fail", 2019. [Online]. Available: https://www.cio.com/article/3429177/6-reasons-why-ai-projects-fail.html.

[4] Pactera, "Pactera white paper reveals 85 percent of AI projects ultimately fail", 2019. [Online]. Available: https://en.pactera.com/news/pactera-white-paper-reveals-85-percent-ai-projects-ultimately-fail/.

[5] A. Chen, "IBM's Watson gave unsafe recommendations for treating cancer", 2018. [Online]. Available: https://www.theverge.com/2018/7/26/17619382/ibms-watson-cancer-ai-healthcare-science.

[6] F. F. Li, "How to Make A.I. That's Good for People", Ney York Times, March 2018. [Online]. Available: https://www.nytimes.com/2018/03/07/opinion/artificial-intelligence-human.html.

[7] d. H. P. I. o. D. a. Stanford, "An Introduction to Design Thinking", [Online]. Available: https://dschool-old.stanford.edu/sandbox/groups/designresources/wiki/36873/attachments/74b3d/ModeGuideBOOTCAMP2010L.pdf.

[8] E. Ries, The Lean Startup: How Today's Entrepreneurs Use Continuous Innovation to Create Radically Successful Businesses, New York: Crown Publishing Group, 2017.

Chapter Technology 4.2

[1] F. M. Roth. An interoperability framework for pervasive computing systems. Mannheim [PhD Thesis], 2018.

[2] F. M. Roth, C. Becker, G. Vega, and P. Lalanda. XWARE - A customizable interoperability framework for pervasive computing systems. Pervasive and Mobile Computing, 47:13-30, 2018.

[3] F. M. Roth, M. Pfannemüller, C. Becker, and P. Lalanda. An Interoperable Notification Service for Pervasive Computing. In Proceedings of International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), pages 842-847. IEEE, 2018.

[4] F. M. Roth, C. Krupitzer, S. VanSyckel, and C. Becker. Nature-Inspired Interference Management in Smart Peer Groups. In Proceedings of International Conference on Intelligent Environments (IE), pages 132-139. IEEE, 2014.

[5] F. Buschmann, R. Meunier, H. Rohnert, P. Sommerlad, and M. Stal. Pattern-Oriented Software Architecture-Volume 1: A System of Patterns. 1996.

Chapter Technology 4.3

[1] A. Vellido, P.J. Lisboa, J. Vaughan. NNs in business: a survey of applications (1992–1998). Expert Systems with applications. 1999

[2] D. Silver, Mastering the game of Go with deep neural networks and tree search. Nature. 2016

[3] Doyle-Lindrud S. Watson will see you now: a supercomputer to help clinicians make informed treatment decisions. Clin J Oncol Nurs 2015;

[4] M.A. Nielsen. NNs and deep learning. San Francisco, CA, USA:: Determination press; 2015

[5] G.E. Hinton. Learning multiple layers of representation. Trends in cognitive sciences. 2007

[6] X.W. Chen, X. Lin. Big data deep learning: challenges and perspectives. IEEE access. 2014

[7] H. Larochelle, Y. Bengio, J. Louradour , P. Lamblin. Exploring strategies for training deep NNs. Journal of machine learning research. 2009

[8] A. Labrinidis, J.V. Hosagrahar. Challenges and opportunities with big data. Proceedings of the VLDB Endowment. 2012

[9] S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber. Gradient flow in recurrent nets: the difficulty of learning long-term dependencies.
[10] Montavon G, Samek W, Müller KR. Methods for interpreting and understanding deep NNs. Digital Signal Processing. 2018

[11] A.R. Ahmad, C. Viard-Gaudin, M. Khalid, E. Poisson. Comparison of support vector machine and neural network in character level discriminant training for online word

recognition. UNITEN Students Con on Research and Development, Malaisie. 2004 Dec 9.

[12] C. Jacobsen, U. Zscherpel, P. Perner. A comparison between NNs and decision trees. International Workshop on Machine Learning and Data Mining in Pattern Recognition.1999

[13] G. Hepner, T. Logan, N. Ritter, N. Bryant. Artificial neural network classification using a minimal training set- Comparison to conventional supervised classification. Photogrammetric Engineering and Remote Sensing. 1990

[14] M.L. Antonie, O.R. Zaiane, A. Coman. Application of data mining techniques for medical image classification. Proceedings of the Second International Conference on Multimedia Data Mining 2001

[15] D. Itchhaporia, P.B. Snow, R.J. Almassy, W.J. Oetgen. Artificial NNs: current status in cardiovascular medicine. Journal of the American College of Cardiology. 1996

[16] P.J. Shayler, M. Goodman, T. Ma. The exploitation of NNs in automotive engine management systems. Engineering Applications of Artificial Intelligence. 2000

[17] I.V. Serban, A. Sordoni, Y. Bengio, A. Courville, J. Pineau. Building end-to-end dialogue systems using generative hierarchical neural network models. Thirtieth AAAI Conference on Artificial Intelligence 2016

Chapter Technology 4.5

[1] Rusnjak, Andreas (2018): Customer Experience im Zeitalter des Kunden. Best Practices, Lessons Learned und Forschungsergebnisse, Wiesbaden.

[2] Gentsch, Peter (2018): Künstliche Intelligenz für Sales, Marketing und Service. Mit AI und Bots zu einem Algorithmic Business - Konzepte, Technologien und Best Practice, Wiesbaden.

[3] Messina, Chris (2016): n.T., URL: https://medium.com/chris-messina/2016-will-be-the-year-of-conversational-commerce-1586e85e3991, 19. September 2018.
[4] Statista (2019): Size of the chatbot market worldwide, in 2016 and 2025 (in million U.S. dollars), URL: https://www.statista.com/statistics/656596/worldwide-chatbot-market/, 13.11.2019

[5] Gartner (2018): Market Guide for Conversational Platforms.

[6] BOTfriends GmbH (2019): Conversational Platform, URL: https://botfriends.de/, 13. October 2019.

[7] MHP Management- und IT-Beratung GmbH (2018): Online Car Sales Studie 2018, Chancen und Handlungsempfehlungen für den Automobilvertrieb via Internet.

[8] Statista (2018): Conversational Commerce Statista Dossier, URL: https://de.statista.com/statistik/studie/id/41378/dokument/conversational-commerce-statista-dossier/

[9] Brandstätter, Ralf (2018), URL: https://www.welt.de/wirtschaft/article185131950/VW-nimmt-seinen-Kunden die-grosse-Auswahl-weg.html, 11.11.2019

Chapter Responsibility 5.1

[1] Pichai, S. (2018). AI at Google: our principles. Retrieved from https://www.blog.google/technology/ai/ai-principles/

[2] Microsoft AI Principals. Retrieved from https://www.microsoft.com/en-us/ai/our-approach-to-ai

[3] Trusted AI for Business. Retrieved from https://www.ibm.com/watson/ai-ethics/

[4] Manyika, J., Chui, M., Joshi, R. (2018). Modeling the global economic impact of AI. McKinsey Global Institute

[5] Universal Declaration of Human Rights. Retrieved from https://www.un.org/en/universal-declaration-human-rights/

[6] Ehlers-Davidsen, J. (2019, August 24). Personal interview

[7] Granville, K. (2018). Facebook and Cambridge Analytica: What You Need to Know as Fallout Widens. Retrieved from https://www.nytimes.com/2018/03/19/technology/facebook-cambridge-analytica-explained.html

[8] Agrawal, N., Goldfarb, A., Gans, J. (2018). Prediction machines: The simple economics of artificial intelligence. Harvard Business Review Press, Brighton, MA.

Chapter Responsibility 5.2

[1] United Nations Climate Change (2019): António Guterres: Climate Change Is Biggest Threat to Global Economy; https://unfccc.int/news/antonio-guterres-climate-change-is-biggest-threat-to-global-economy

[2] United Nations: Sustainable Development Goals; https://www.un.org/sustainabledevelopment/sustainable-development-goals/

[3] United Nations, "Transforming our World: The 2030 Agenda for Sustainable Development", 2015.

[4] Umwelt-Bundesamt (2018): Treibhausgasemissionen in Deutschland 1990 bis 2018 in Mio t CO2-äquivalent; https://www.umweltbundesamt.de/sites/default/files/medien/2294/bilder/thg_nach_gasen_0.jpg

[5] United States Environmental Protection Agency: Greenhouse Gas Emissions; https://www.epa.gov/ghgemissions/sources-greenhouse-gas-emissions

[6] Forbes (2019): How AI Is Making Buildings Smart And Intelligent; https://www.forbes.com/sites/cognitiveworld/2019/08/13/how-ai-is-making-buildings-smart-and-intelligent/#722a68a528d7

[7] World Health Organization (2018): Global hunger continues to rise, new UN report says; https://www.who.int/news-room/detail/11-09-2018-global-hunger-continues-to-rise---new-un-report-says

[8] Focus Economics (2018): The Poorest Countries in the World; https://www.focus-economics.com/blog/the-poorest-countries-in-the-world

[9] Bill and Melinda Gates Foundation, "Goalkeeper: The Stories Behind the Data 2018", 2018.

[10] United Nations (2017): World population projected to reach 9.8 billion in 2050, and 11.2 billion in 2100; https://www.un.org/development/desa/en/news/population/world-population-prospects-2017.html

[11] World Health Organization: Drinking-water; https://www.who.int/news-room/fact-sheets/detail/drinking-water

[12] U.S. Geological Survey (2018): Contamination of Groundwater; https://www.usgs.gov/special-topic/water-science-school/science/contamination-groundwater?qt-science_center_objects=0#qt-science_center_objects

[13] Allerin (2018): 4 ways AI can revolutionize waste management; https://www.allerin.com/blog/4-ways-ai-can-revolutionize-waste-management

[14] United States Environmental Protection Agency: Sustainable Manufacturing; https://www.epa.gov/sustainability/sustainable-manufacturing

[15] Vialytics; https://vialytics.de/

[16] Deepmind (2016): DeepMind AI Reduces Google Data Centre Cooling Bill by 40%; https://deepmind.com/blog/article/deepmind-ai-reduces-google-data-centre-cooling-bill-40

[17] ESRI (2018): Geography and AI Combine for Business Intelligence; https://www.esri.com/about/newsroom/publications/wherenext/artificial-intelligence-in-business/

[18] Tomra; https://www.tomra.com/en

[19] McKinsey (2017): Smartening up with Artificial Intelligence (AI) -What's in it for Germany and its Industrial Sector?; https://www.mckinsey.com/~/media/McKinsey/Industries/Semiconductors/Our%20Insights/Smartening%20up%20with%20artificial%20intelligence/Smartening-up-with-artificial-intelligence.ashx%20str%209

[20] WINT; https://wint.ai

[21] Intuitive AI; https://intuitiveai.ca/

[22] Business Insider (2018): The chairman of the company behind Snickers and M&Ms reveals what's next for the chocolate giant; https://www.businessinsider.de/mars-chairman-stephen-badger-reveals-details-about-company-future-2018-5?r=US&IR=T

CREDITS

Photo credits

Seite 1 ©shutterstock – Rawpixel.com
Seite 6-7 ©shutterstock – guruXOX
Seite 14-15 ©shutterstock – Pormezz
Seite 18 ©unsplash – Andy Beales
Seite 22-23 ©gettyimages – martin-dm
Seite 27 ©shutterstock – spawn101
Seite 29 ©gettyimages – Eric Larrayadieu
Seite 32-33 ©gettyimages – Adam Gault
Seite 36-37 ©shutterstock – ra2studio
Seite 41 ©unsplash – Jordan Harrison
Seite 51-52 ©unsplash – Alexandre Debieve
Seite 74 ©shutterstock – MintArt
Seite 78 ©unsplash – Evgeni-evgeniev
Seite 80-81 ©shutterstock – MIND AND I
Seite 86-87 ©shutterstock – mimagephotography
Seite 94 ©gettyimages – cicerocastro

THE AUTHORS

BENEDIKT BOTHUR

Service Unit
Intelligent Software

Tasks@MHP
Machine Learning Engineer.

A.I.dea Book Role
Author

Professional Interests
Intersection of sustainability and Artificial Intelligence
Machine Learning with a focus on Deep Learning
Scalability of AI models and their integration into complex infrastructures

YANNICK FISCHER

Service Unit
Intelligent Software

Tasks@MHP
Technical Consultant with a focus on Machine Learning and project development. In-vehicle development of AI functions.

A.I.dea Book Role
Author

Professional Interests
Machine Learning
Sustainable implications of AI
Global economic development
AI in financial markets

DR. WILLIAM COBBAH

Service Unit
Intelligent Software

Tasks@MHP
Artificial Intelligence Business Developer.
Project Manager & Solution Architect.

A.I.dea Book Role
Editor and Author

Professional Interests
Economic, environmental and social impact of Artificial Intelligence
AI applications in healthcare and neuroscience
Cloud-based application architectures
Machine Behavior

STEFFEN FRANZ

Service Unit
Digital Supply Chain Solutions

Tasks@MHP
Solution Architect and expert for Supply Chain Planning.
Head of Business Area Advanced Business Planning.

A.I.dea Book Role
Author

Professional Interests
Supply Chain Planning Consulting and Project Management
Enhancing integrated supply chain planning capabilities with end-to-end value-chain visibility
Integration of Machine Learning and optimization algorithms into Supply Chain Planning and integrated business planning

ARNO FUHRMANN

Service Unit
Intelligent Software

Tasks@MHP
Data Scientist and Full Stack Developer with experience in interdisciplinary projects in electrical, mechanical and aerospace engineering.

A.I.dea Book Role
Author

Professional Interests
Computer Vision, Image Recognition, Deep Learning, and Cognitive Systems
Machine Learning, Decision Making and Problem Solving
Economic, environmental and social impact of Artificial Intelligence

MICHAEL GENNE

Service Unit
Big Data Analytics & IoT Technologies

Tasks@MHP
Data Strategy Consultant with focus on conception and implementation of data platforms as well as the operationalization of data governance frameworks.

A.I.dea Book Role
Author

Professional Interests
Artificial Intelligence use cases from various industries
Agile collaboration models and project management methods

DOMINIK GRAETZ

Service Unit
Big Data & IoT Technologies

Tasks@MHP
AI & data strategist with a focus on use case management & data governance. Guiding the AI & data-driven transformation from a strategical, organizational & technological perspective.

A.I.dea Book Role
Author

Professional Interests
Transforming data into a primary asset
Data thinking & data driven business models
Value driven data governance
Analytics enabled product & service development

BEATRICHE GENCHEVA

Service Unit
Big Data & IoT Technologies

Tasks@MHP
Integration of CarOnDemand Services. SAP System Integration. Development of IoT-Prototypes. Software Engineer with a focus on Machine Learning.

A.I.dea Book Role
Author

Professional Interests
IoT Cloud Platforms
Fraud Detection
Social impact of Artificial Intelligence in Online Social Networks

DR. TILL GIESE

Company
Porsche Consulting

Service Unit
Industrial Goods – Smart Operations

Tasks@Porsche Consulting
Senior Manager with focus on Operational Excellence, Efficiency Programs, and Industry 4.0.

A.I.dea Book Role
Author

Professional Interests
Business applications for AI
Big Data analysis
Corporate wide efficiency programs
Global business systems

SEBASTIAN GROSSE

Service Unit
Mobility Services & Strategy

Tasks@MHP
UX Engineer with a focus on user centricity and service design in the various divisions within the automotive industry.

A.I.dea Book Role
Author

Professional Interests
User Experience Design
Information Design
Rapid Prototyping & Testing
Future Thinking

ALEXANDER HÖWELER

Service Unit

Tasks@MHP
Data Scientist focusing on Customer Insight Analysis with a combined background of business and analytical knowledge.

A.I.dea Book Role
Author

Professional Interests
Creating significant and sustainable business impact through Data Science Projects
Natural Language Processing with focus on Customer Interaction Analysis
Custom-tailored visualization of analytical results

JAN JANETZKO

Service Unit
Digital Supply Chain

Tasks@MHP
Technical Consultant with a focus on Digital Supply Chain Solutions specifically Digital Manufacturing and MES.

A.I.dea Book Role
Author

Professional Interests
Shaping the future of Industrial Engineering and Production Management
AI enhanced Quality Management
Machine integration and IoT
MES implementation and integration

JOHANNES KELLER

Service Unit
Intelligent Software

Tasks@MHP
Head of Business Area Artificial Intelligence, Portfolio Developer Artificial Intelligence, Program/Project Manager.

A.I.dea Book Role
Editor

Professional Interests
Data driven products
Large-scale Machine Learning Architectures
Agile organization in governance and projects

FANLI LIN

Service Unit
Intelligent Software

Tasks@MHP
Full-Stack Data Scientist with a focus on Computer Vision and Machine Learning

A.I.dea Book Role
Author

Professional Interests
Designing and building robust production-ready machine learning models
Leveraging statistical techniques to answer business questions using available data
Implementing interactive data visualization dashboards
Big Data engineering

ALMAS MYRZATAY

Service Unit
Data Science, MHP Americas

Tasks@MHP
Consultant and Software Developer with a focus on Machine Learning and Big Data; Implementation of CI/CD pipeline and System Architecture designs.

A.I.dea Book Role
Author

Professional Interests
Deep Learning with focus on Computer Vision
Applications of machine learning in Internet of Things (IoT) devices
Democratization of AI

POOJA MUKUND

Service Unit
Data Science, MHP Americas

Tasks@MHP
Data Scientist

A.I.dea Book Role
Author

Professional Interests
Statistical Modeling and Machine Learning
AI applications in healthcare and neuroscience
Ethics of AI

HOLGER MUTH-HELLEBRANDT

Service Unit
Intelligent Software

Tasks@MHP
Computational Linguist with focus on Machine Learning, Big Data and Realization of End-to-End AI Solutions.

A.I.dea Book Role
Author

Professional Interests
Challenges arising from turning AI PoCs into Enterprise Solutions and Software in Production
Deep Learning applied to Computer Vision
Natural Language Processing
AI Solutions as Enablers

CHRISTOPH NABER

Service Unit
Intelligent Software

Tasks@MHP
Software Architect and Engineer. Backend design and development in various environments.

A.I.dea Book Role
Author

Professional Interests
AI Design patterns
Genetic algorithms
Process design
XSLT development

DR. FELIX MAXIMILIAN ROTH

Service Unit
Intelligent Software

Tasks@MHP
Software and Cloud Engineer with a focus on Internet of Things and Artificial Intelligence.

A.I.dea Book Role
Author

Professional Interests
Internet of Things, Machine Learning, Artificial Intelligence, Cloud and Edge Computing
Sustainability through Internet of Things and Artificial Intelligence
Healthcare and personal health through Internet of Things and Artificial Intelligence

KATARINA PREIKSCHAT

Service Unit
Intelligent Mobility

Tasks@MHP
Product Development with a focus on Technical Design and User Experience. Interface of Users, Designers and Developers.

A.I.dea Book Role
Author

Professional Interests
Project Management
Data Exploration and Data Science (especially Machine and Deep Learning)
Problem Identification and Solution Design
Distributed Ledger Technologies (Consulting and Development)

HEATHER A. SMITH

Service Unit
Data Science, MHP Americas

Tasks@MHP
Head of Data Science, building MHP Americas' data science and AI/ML program.

A.I.dea Book Role
Author

Professional Interests
Machine Learning, Deep Learning applied to Computer Vision and Natural Language Processing
AI robotics
Social and economic impact of Artificial Intelligence
Ethical issues arising from Artificial Intelligence

JOHANNES STARK

Service Unit
Big Data & IoT Technologies

Tasks@MHP
Disciplines and methods in the field of Data Science with focus on Statistical Modeling and Predictive Analytics. Designing and programming Data Analytics pipelines to generate added value from raw data.

A.I.dea Book Role
Author

Professional Interests
Linking statistical and Machine Learning approaches
Analyzing unstructured data with Natural Language Processing

THILLAI SIVAKUMARAN

Service Unit
Mobility Services & Strategy

Tasks@MHP
Project Management with a Focus on end-to-end delivery in innovation projects.

A.I.dea Book Role
Author

Professional Interests
Data Thinking
Data Exploration and Data Science (especially Machine and Deep Learning)
Problem Identification and Solution Design
Human-centered Artificial Intelligence in the Context of Business Process Optimization

VANESSA VIELLIEBER

Service Unit
Big Data & IoT Technologies

Tasks@MHP
Senior Data Scientist, understanding customer's challenges, developing target-oriented solutions and leading them to success

A.I.dea Book Role
Author

Professional Interests
Natural Language Processing using statistics, mathematics, linguistics, and deep learning
Research in Deep Learning and Knowledge Graphs

NIKOLAJ WALLER

Service Unit
Big Data & IoT Technologies

Tasks@MHP
AI Engineer, Data Scientist and Software Developer

A.I.dea Book Role
Author

Professional Interests
Computer Vision
Natural Language Processing
Machine Learning and its application
Software Development for AI Use Cases

FABIAN WITTKE

Service Unit
Intelligent Software

Tasks@MHP
Software Engineering and architecture with a focus on Machine Learning and Big Data processing. End-to-end integration of complex cloud systems.

A.I.dea Book Role
Author

Professional Interests
Production ready Machine Learning Systems
Architecture and design of complex software systems
Value creation with Artificial Intelligence

ABOUT MHP

MHP is a globally active and leading management and IT consultancy. We develop groundbreaking mobility and manufacturing solutions for international corporations, established medium-sized companies and disruptive start-ups. As a premium business and technology partner, we are today shaping the digital future of tomorrow.

Our consulting approach is unique: We combine holistic IT and technology expertise with in-depth management know-how. This makes MHP the ideal partner for a successful digital turnaround. As digitalization experts, we provide innovative strategies based on well-founded analyses to transform change processes into lasting success.

With more than 2,500 employees at 12 locations around the world, we drive digital progress – in collaboration with more than 300 clients and with excellence at all levels.

Contact us

If you would like to know about our approach to Artificial Intelligence you can contact:

Johannes.Keller@mhp.com
William.Cobbah@mhp.com

MHP: DRIVEN BY EXCELLENCE

12 **MHP**Offices in Germany, United Kingdom, USA, China and Romania

www.mhp.com